Automatic Detection for Dam Restored Concrete Based on DeepLabv3+

Zhang Jun*, Chen Shiqiao², Deng Zhenhua², Zeng yuan², Li Chabei²

¹State Grid Hunan Electric Power Company Limited Research Institute, Changsha 410007, China
²Hydropower Branch of State Grid Hunan Electric Power Company Limited, Changsha 410000, China
*2011102060033@whu.edu.cn
*Corresponding author’s e-mail: sghnzhj1988@163.com

Abstract: The concrete surface defects caused by various factors need to be repaired since the construction stage of the concrete dam. However, the binding strength between restored concrete and the original concrete will gradually get worse in the long run and appear some problems such as peeling. At present, the repaired concrete is mainly checked by the workforce, but it is time-consuming, inefficient, and hard to quantitative evaluate, such as the peeling area. A semantic segmentation method based on the DeepLabv3+ with ResNet50 backbone is proposed to restored concrete identification automatically. The dam restored concrete data set is established, including 372 high-resolution images to verify the method. The results indicated that the DeepLabv3+ model finally reaches 0.68 mIoU on the test set, which is a practical way to detect dam restored concrete.

1. Introduction
Dam concrete belongs to the large volume concrete, and its external appearance is wide. At the dam construction stage, the concrete surface will inevitably appear some defects such as the wrong table, hang curtain, the leakage steel heads, and pipe fittings, pits, bubbles, honeycomb, quality defects [1-2]. No matter the inevitable objective reason or subjective human factors caused, these concrete defects require appropriate and necessary repair to meet the overall quality and appearance requirements. Therefore, it is vital to repair the concrete surface defect for safe operation and the dam's service life.

However, the restored concrete is not as strong as the original concrete. The binding strength between repaired concrete and the original concrete will gradually get worse in the long run and appear the phenomenon such as peeling. Manual inspection is a commonly used monitoring method for dam restored concrete at present. However, there are some problems such as time-consuming, difficulty in comprehensive coverage, and low efficiency.

In recent years, the deep learning technique represented by the Fully convolutional networks (FCN) [3] is continuously surpassing the previous image semantic segmentation techniques in the image pixel-level classification task. It is gradually being applied in various industries.

In this paper, on-site restored concrete image data of a concrete dam is collected. A restored concrete data set is established, and then is identified by using the DeepLabv3+ technology based on FCN, thus providing an effective way to automatic detection and liberate human resources.
2. DeepLabv3+
DeepLab series [4-7] are semantic segmentation deep learning models developed from FCN model. There are two technical hurdles in the application of basic FCN model: downsampling, and spatial invariance. The downsampling method will reduce the resolution, especially at the high-level layers. The spatial invariance means that obtaining object-centric decisions from a classifier requires invariance to spatial transformations, inherently limiting the spatial accuracy. DeepLabv1 employs the atrous convolution algorithm and conditional random field (CRF) to address the down-sampling and the spatial invariance, respectively. DeepLabv2 uses atrous spatial pyramid pooling (ASPP) to robustly segment objects at multiple scales. To encode multi-scale information, DeepLabv3 proposes a cascaded module and an improved ASPP module. The cascaded module gradually doubles the atrous rates and the improved ASPP module augmented with image-level features probes the features with filters at multiple sampling rates and effective field-of-views. DeepLabv3+ extends DeepLabv3 by adding a simple yet effective decoder module to refine the segmentation results, especially along object boundaries. For backbone, DeepLabv1 is constructed by VGG-16. DeepLabv2 and DeepLabv3 use the ResNet. DeepLabv3+ (see Figure 1) adopts the ResNet and Xception.

3. The model flow chart
The flowchart of our semantic segmentation method is shown in Figure 2 and demonstrated as follows:
- S1: Obtain images from the concrete surface of the dam, and label the images with or without restored concrete pixel by pixel;
- S2: The image and its labelled image are processed by size scaling;
- S3: The generated scaled images are divided into a training set and test set, which are used for model training and test;
- S4: Train the DeepLabv3+ according to the training set in step S3;
- S5: The model trained according to step S4 is used to classify the images pixel by pixel in the test set from step S3.
4. Examples
The image data is derived from the restored concrete of an arch dam, which is obtained by taking photos with a mobile phone camera. The resolution of the image of restored concrete is 3456*4608, and a total of 372 on-site images are collected. The images are scaled to 513*513. According to the ratio of 4:1, the scaled images were divided into a training set and test set, respectively. The DeepLabv3+ and the backbone ResNet50 are adopted. The deep learning library Tensorflow is used for model training and prediction.

The cross-entropy loss function is used as follows:

\[
\text{Loss} = -y \log y' - (1 - y) \log (1 - y')
\]

where, \(y\) denotes the positive and negative label; \(y'\) denotes the output probability.

The evaluation indexes of the model are as follows:

\[
\text{mIoU} = \frac{1}{M} \sum_{m=1}^{M} \frac{TP}{TP + FP + FN}
\]

where, \(M\) denotes the images' number. \(TP\) denotes the true positive predictions; \(FP\) means false positive predictions; \(FN\) denotes the false-negative predictions.

The model training process is as follows:

![Training process of DeepLabv3+](image-url)
For the Deeplabv3+, the initial learning rates are set as 0.001, and the Adam updater is used. The weight decay to use for regularizing the model is set $10^{-4}$ to prevent overfitting. The Batch Normalization (BN) parameter is set $\epsilon=10^{-6}$, which is a small constant to prevent division by zero when normalizing activations by their variance in BN. The Deeplabv3+ are trained for 1000 epochs, and the batch size is set as 4.

As can be seen from Figure 3, the Loss and mIoU converge after about 200 epochs on both the training and test sets except the training Loss, which decreases gradually. However, there are some unstable jump points in the training process. The mIoU can achieve 0.68 on the test set. Meanwhile, Figure 4 shows some of the semantic segmentation results on the test set, which indicates that the Deeplabv3+ is a powerful method to detect the restored concrete of the dam.
5. Conclusion
This study adopts a segmentation method for restored concrete detection. The Deeplabv3+ is developed from the classic FCN and can achieve pixel-level prediction. The details through the pixel-level information, such as the area of restored concrete, can be acquired automatically, which is very suitable for restored concrete estimation. The on-site dam restored concrete images are used to verify the Deeplabv3+, which is finally proved to be useful.

Acknowledgments
This work is supported by the State Grid Hunan Electric Power Company Limited science project (No. 5216A518000N). The authors would like to thank the Hydropower Branch of State Grid Hunan Electric Power Company Limited for providing the verification data.

References
[1] Zhu Yongbin, Cao Dengyun, Ren Yinxia. (2019) Repair and protection of concrete defects in dam concrete panels. New building materials. 46(09):94-97+122.
[2] Wang Da-Zhang. Repair of concrete surface defects in Ertan Dam. Design of Hydropower Station, 02:93-97. 1998.
[3] Long, J., Shelhamer, E., & Darrell, T. (2015) Fully convolutional networks for semantic segmentation. Computer vision and pattern recognition. 3431-3440.
[4] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014) Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs. arXiv: Computer Vision and Pattern Recognition.
[5] Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018) DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 834-848.
[6] Chen, L., Papandreou, G., Schroff, F., & Adam, H. (2017) Rethinking Atrous Convolution for Semantic Image Segmentation. arXiv: Computer Vision and Pattern Recognition.
[7] Chen, L., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. European conference on computer vision.