Recognition and Application of Tunnel Water Accumulation Based on Computer Vision

Yuan Chen1*, Yongwei Wang1,2 and Kunyao Li1
1 CCCC Second Harbor Engineering Company LTD, Wuhan, Hubei, 430040, China
2 Key Laboratory of Large-span Bridge Construction Technology, Wuhan, Hubei, 430040, China
*Corresponding author’s e-mail: chenyuan20@ccccltd.cn

Abstract. Water accumulation in tunnel threatens the safety of driving and the tunnel itself. In order to detect the tunnel waterlogging in time, a method based on computer vision is proposed. This method utilizes Laplace transform for image preprocessing to remove fuzzy image, uses the network MobileNetV2 to build a tunnel waterlogging recognition model, and then smooths the prediction results. Based on this method, a tunnel waterlogging recognition and early warning platform is developed which applied to a city video surveillance system successfully. The results show that: the tunnel waterlogging recognition method based on computer vision has high accuracy, tiny computation and low cost. The existing video monitoring system can be upgraded intelligently without installing any hardware, so as to realize the active recognition and early warning of tunnel water accumulation.

1. Introduction

Tunnels are important part of the national transportation infrastructure. At the end of 2020, there are 21,316 road tunnels nationwide, including 1,394 extra-long tunnels[1]. Due to frequent and sudden extreme weather, the phenomenon of water accumulation in tunnels is increasing year by year. Water accumulation in tunnels not only affects the passage of vehicles and causes traffic jam, but also threatens the safety of driving and the tunnel itself. Therefore, it is of great significance to quickly identify the accumulation of water in tunnel and to give early warning in real time.

At present, tunnel waterlogging monitoring is mainly obtained through various sensors[2]. With the development of big data, data-driven recognition methods have begun to gain attention. This method comprehensively considers weather, rainfall, traffic flow, vehicle speed, historical water accumulation to predict road water accumulation[3]. In recent years, with the rapid development of artificial intelligence technology, non-contact recognition methods based on computer vision have gradually been applied. Omer et al extract RGB features along with gradients to train a traditional machine learning model Support Vector Machine (SVM) which is used to classify the images into their respective categories[4]. In order to distinguish road surface condition at night-time, Takeuchi et al exploit illuminated object such as street lights, signal lights, reflections and other lighting sources to extract the image features, and then classify road surface conditions at night-time, such as dry, wet and snow by these features[5].

Tunnel waterlogging recognition is a binary classification problem. The image recognition method based on deep learning has shown extremely high performance on classification tasks. In 2012, Hinton et al designed the AlexNet network, which achieved a 16% error rate in the ImageNet competition,
exceeding the traditional image processing algorithm[6]; in 2015, He et al developed ResNet with a depth of 152 layers and achieved an error rate of 4%, and the recognition accuracy rate surpassed that of humans for the first time[7]; in 2017, SENet proposed by Hu et al reduced the error rate to 2.25%[8]. In order to improve the accuracy and speed of model at the same time, more and more efficient models have been proposed one after another, such as ShuffleNet[9], MobileNet series[10][11].

Although a small amount of research has been done on the recognition method of tunnel waterlogging based on convolutional neural network, its recognition speed and accuracy need to be further improved, and the related application research is still deficient. In this paper, an efficient neural network MobileNetV2 is used to construct a tunnel waterlogging recognition model where the Laplace transform is applied for image pre-processing. Finally, a low-cost tunnel waterlogging recognition and early warning system was developed.

2. Image classification of tunnel waterlogging
For a given image, classification task uses a certain algorithm to determine the category of the image. Supervised learning refers to the process of using a set of samples of known categories to adjust the parameters of the classifier to achieve the required performance. MobileNetV2 is an efficient mobile network developed by Google. It has faster speed and relatively high accuracy and is widely used in engineering projects.

In this paper, the mobileNetV2 network is used for tunnel waterlogging image classification and supervised learning methods is used for model training. A binary classification model is trained using the tunnel waterlogging dataset. In order to adapt to our task, the final fully connected layer of the model is adjusted.

2.1. Tunnel waterlogging image dataset
A total of 3,600 images of tunnel waterlogging were obtained through video surveillance and web crawler, of which 80% were used for training set and 20% were used for test set. The data is manually labelled, including "waterlogging" and "normal" (Figure 1). In order to improve the robustness of the model, random cutting, translation, rotation, flipping, brightness adjustment, and contrast are used for data augmentation.

The evaluation indicators for classification problems include: accuracy, precision, recall and $F_1$ score. Among them, $F_1$ score is the harmonic mean of precision and recall, which can punish extreme situations and is mostly used to measure the comprehensive ability of the model. In this paper, $F_1$ score is selected for model evaluation.

$$ precision = \frac{TP}{TP + FP} \quad (1) $$
$$ recall = \frac{TP}{TP + FN} \quad (2) $$
$$ F_1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (3) $$

![Figure 1. Example of tunnel waterlogging dataset.](image-url)
2.2. Image preprocessing
On-site existing video surveillance images are transmitted to the background waterlogging recognition and early warning system through the network. Due to the unstable network, the image may have fuzzy phenomenon and cause the model to misjudge, as shown in Figure 2. Therefore, it is necessary to pre-process the returned live images.

![Fuzzy and clear images](image)

Figure 2. Example of fuzzy and clear image.

The Laplace transform is mostly used to solve the edge of the image, and its essence is to calculate the second derivative of the image to find the area with rapid density changes. In a normal image, the boundary is relatively clear, and the variance of the second-order derivative of the image will be relatively large, while the fuzzy image is relatively confused and contains fewer boundaries, so the variance will be relatively small. Based on this feature, the fuzzy image can be recognized and removed.

The Laplacian is an isotropic differential operator with rotation invariance, which is defined as:

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

(4)

In a two-dimensional digital image, the second-order differences in the two directions are:

$$\frac{\partial^2 f}{\partial x^2} = f(x+1, y) + f(x-1, y) - 2f(x, y)$$

(5)

$$\frac{\partial^2 f}{\partial y^2} = f(x, y+1) + f(x, y-1) - 2f(x, y)$$

(6)

Substituting equations (5) and (6) into (4), the Laplace transform of the image obtained is:

$$\nabla^2 f(x, y) = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y)$$

(7)

The fuzzy image recognition process based on Laplace transform is as follows:

1. Image grayscale: Convert the image from a 3-channel colourful image to a single-channel black and white image;
2. Image scaling: adjust the image to a uniform size, which is convenient for setting the judgment threshold;
3. Image Laplace transform: Perform Laplace transform on the image generated in (2) above to obtain the edge image;
4. Calculate the variance of the image after Laplace transform: solve the variance of the image generated by the above (3), set a unified judgment threshold, if it is less than the threshold, the image may be blurred which can be removed directly. Otherwise, the image is normal.

2.3. Prediction result smoothing
When using the image binary classification recognition model to identify the returned live video surveillance image, if the judgment is just based on the single-frame, a sort of “prediction flickering” will be encountered due to vehicle lights, water surface fluctuations, etc., the recognition result will frequent jump between "waterlogging" and "normal". In order to suppress this phenomenon, continuous $N$ frames of model recognition results are used and the average value is used for smoothing. The calculation formula is as follows:
\[ P = \frac{\sum_{i=1}^{N} P_i}{N} \] (8)

Where \( P_i \) is the probability of \( i \)-th frame whose value is between 0 and 1. In order to improve the robustness of the recognition result and reduce misjudgement, the alarm is only given when the recognition result of consecutive \( M \) frames is "waterlogging".

3. Model training and system construction

3.1. Model training and evaluation

The following deep learning model is implemented in the PaddlePaddle framework with NVIDIA’s CUDA for neural network acceleration. The hardware configuration is as follows: Intel i9-9900@3.10GHz with 64GB RAM and Nvidia Quadro P4000 8G. The learning rate is 0.01 with a batch size of 64 examples, and the Stochastic Gradient Descent (SGD) method is used to solve the model. There are 100 epochs of training which takes 35 minutes totally.

The loss function is a measure of inconsistency between the model’s prediction and true labels. The loss curve of the training process is shown in Figure 3. In the first 50 Epochs of training, the loss of the model continues to decrease, which indicates that the initial model has a poor ability to recognize tunnel waterlogging with a larger loss. With learning continues, the loss value is getting smaller and smaller, and the model recognition capabilities continue to increase. In the last 50 epochs of training, the loss value tends to be stable, indicating that the model has normally converged and training can be ended.

![Figure 3. MobileNetV2 model training curve.](image)

In order to evaluate the performance of the model quantitatively, the aforementioned \( F1 \) value is used to evaluate the model effect, and four typical classification models are compared, as shown in Table 1. The model inference time is the average of 100 tests on the Nvidia Quadro P4000 8G as the final, not including the image reading time. As can be seen from the table, the MobileNetV2 model has achieved a good balance between accuracy and time. The \( F1 \) value of the model reached 0.944 and the inference time was only 21 milliseconds. Compared with the most accurate model Xception, although the accuracy of MobileNetV2 was reduced by about 2%, the reasoning time is reduced by nearly 1 time, and it is convenient for mobile terminal deployment.
Table 1. Performance comparison of different models on waterlogging dataset.

| Model Name   | Model Size (M) | Inference Time (ms) | F1-score |
|--------------|----------------|---------------------|----------|
| AlexNet      | 232            | 12                  | 0.891    |
| ResNet50     | 98             | 27                  | 0.951    |
| Xception     | 88             | 39                  | 0.963    |
| MobileNetV2  | 14             | 21                  | 0.944    |
| ShuffleNetV2 | 9              | 15                  | 0.927    |

Notes: The inference time uses the average of 100 tests.

3.2. System construction and application
Taking the above-developed binary classification model, a tunnel waterlogging recognition and early warning system is constructed. This system uses the existing video surveillance system on site and can intelligently upgrade the video surveillance system without installing any hardware. The workflow is as follows:

(1) Image data transmission: On-site video monitoring transmits images to the recognition system through frame extraction software;

(2) Image pre-processing: Recognize and remove the fuzzy image that may appear;

(3) Recognition of waterlogging image: Recognize whether tunnel water accumulation occurs based on the deep learning model;

(4) Post-processing of results: smoothing the recognition results of the model to improve the robustness of the recognition results;

(5) Early warning: When water accumulation is identified, early warning information will be released in time.

Figure 4. The overall framework of tunnel waterlogging application.

The summer rainy season is a high incidence of urban waterlogging. In 2020, the system was applied to the urban tunnel monitoring system and upgraded intelligently. The application effect is shown in Figure 5. When there was a sudden rainstorm in the city, water accumulated in a tunnel. The system identified it in time and issued an early warning. After receiving the information, the relevant departments rushed to the site in time for water drainage and traffic diversion. The application results show that the system is stable and reliable, and can meet the needs of engineering applications.
4. Conclusion
In order to identify the tunnel waterlogging in real time, an efficient neural network MobileNetV2 was used to construct a recognition model. The fuzzy image was identified and removed based on the Laplace transform of the image, and a tunnel water accumulation rapid identification method and early warning system based on ordinary video surveillance was developed. The method utilizes the existing video monitoring system at the tunnel site, and can realize intelligent upgrade without installing any hardware with high accuracy, tiny computation and low cost. The system is applied to an urban video surveillance system, and the application results show that the system is stable and reliable and can meet the needs of engineering applications.

References
[1] Ministry of Transport of the People's Republic of China. (2021) Statistics Bulletin of traffic transportation industry Development in 2020. https://xxgk.mot.gov.cn/2020/jigou/zhghs/202105/t20210517_3593412.html.
[2] Zhang J, Wang P, Wang H, et al. (2021) Development and verification of integrated photoelectric system for noncontact detection of pavement ponding and freezing. Structural Control and Health Monitoring, 28(6): e2719.
[3] Tang, X., Qiu, J., & Fu, K. (2020) A Method for Urban Road Waterlogging Recognition Based on Fundamental Diagram. In: International Conference on Transportation and Development. Reston. pp. 231-242.
[4] Omer, R., & Fu, L. (2010) An automatic image recognition system for winter road surface condition classification. In: 13th international IEEE conference on intelligent transportation systems. Madeira Island. pp. 1375-1379.
[5] Kawai, S., Takeuchi, K., Shibata, K., & Horita, Y. (2012) A method to distinguish road surface conditions for car-mounted camera images at night-time. In: 12th International Conference on ITS Telecommunications. Taipei. pp. 668-672.
[6] Krizhevsky A, Sutskever I, Hinton GE. (2012) Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25: 1097-1105.
[7] He K, Zhang X, Ren S, et al. (2016) Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition. Las Vegas. pp. 770-778.
[8] Hu J, Shen L, Sun G. (2018) Squeeze-and-excitation networks. In : IEEE conference on computer vision and pattern recognition. Salt Lake City. pp. 7132-7141.
[9] Zhang X, Zhou X, Lin M, et al. (2018) Shufflenet: An extremely efficient convolutional neural network for mobile devices. In: IEEE conference on computer vision and pattern recognition. Salt Lake City. pp. 6848-6856.
[10] Sandler M, Howard A, Zhu M, et al. (2018) Mobilenetv2: Inverted residuals and linear bottlenecks. In: IEEE conference on computer vision and pattern recognition. Salt Lake City. pp. 4510-4520.
[11] Howard A, Sandler M, Chu G, et al. (2019) Searching for mobilenetv3. In: International Conference on Computer Vision. Thessaloniki. pp. 1314-1324.