Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

A deep learning technique-based automatic monitoring method for experimental urban road inundation

PLEASE CITE THE PUBLISHED VERSION

https://doi.org/10.2166/hydro.2021.156

PUBLISHER

IWA Publishing

VERSION

VoR (Version of Record)

PUBLISHER STATEMENT

This is an Open Access Article. It is published by IWA Publisheing under the Creative Commons Attribution 4.0 Unported Licence (CC BY). Full details of this licence are available at: http://creativecommons.org/licenses/by/4.0/

LICENCE

CC BY 4.0

REPOSITORY RECORD

Han, Hao, Jingming Hou, Ganggang Bai, Bingyao Li, Tian Wang, Xuan Li, Xujun Gao, et al.. 2021. “A Deep Learning Technique-based Automatic Monitoring Method for Experimental Urban Road Inundation”. Loughborough University. https://hdl.handle.net/2134/16681963.v1.

Check for updates

A deep learning technique-based automatic monitoring method for experimental urban road inundation

Hao Han, Jingming Hou, Ganggang Bai, Bingyao Li, Tian Wang, Xuan Li, Xujun Gao, Feng Su, Zhaofeng Wang, Qiuhua Liang and Jiahui Gong

ABSTRACT

Reports indicate that high-cost, insecurity, and difficulty in complex environments hinder the traditional urban road inundation monitoring approach. This work proposed an automatic monitoring method for experimental urban road inundation based on the YOLOv2 deep learning framework. The proposed method is an affordable, secure, with high accuracy rates in urban road inundation evaluation. The automatic detection of experimental urban road inundation was carried out under both dry and wet conditions on roads in the study area with a scale of a few m^2. The validation average accuracy rate of the model was high with 90.1% inundation detection, while its training average accuracy rate was 96.1%. This indicated that the model has effective performance with high detection accuracy and recognition ability. Besides, the inundated water area of the experimental inundation region and the real road inundation region in the images was computed, showing that the relative errors of the measured area and the computed area were less than 20%. The results indicated that the proposed method can provide reliable inundation area evaluation. Therefore, our findings provide an effective guide in the management of urban floods and urban flood-warning, as well as systematic validation data for hydrologic and hydrodynamic models.

Key words | deep learning technique, experimental urban road inundation, inundation area, YOLOv2 model

HIGHLIGHTS

- First experimental urban road inundation automatic detection study using YOLOv2.
- Proposed an inundation area computation method based on a deep learning technique.
- Good performance on an experimental urban road inundation detection was tested.

Hao Han
Jingming Hou (corresponding author)
Ganggang Bai
Bingyao Li
Tian Wang
Xuan Li
Jiahui Gong
State Key Laboratory of Eco-hydraulics in Northwest Arid Region of China,
Xi'an University of Technology,
No. 5 Jinhua Road, Xi' an 710048, Shaanxi,
China
E-mail: *jingming.hou@xaut.edu.cn*

Xujun Gao
Feng Su
Power Construction Corporation of China,
Northwest Engineering Corporation Limited,
Xi' an 710065, Shaanxi,
China

Zhaofeng Wang
Qingyang Municipal Housing and Urban and Rural Construction Bureau,
Qingyang 74500, Gansu,
China

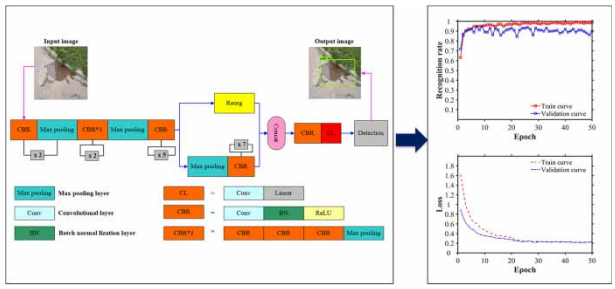
Qiuhua Liang
School of Architecture, Building and Civil Engineering,
Loughborough University,
Loughborough,
UK

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

doi: 10.2166/hydro.2021.156

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

GRAPHICAL ABSTRACT



INTRODUCTION

Rising urbanization has caused increased stormwater runoff in recent years. This phenomenon has been attributed to the transformation of previous semirural environments into urban infrastructure setup (Baek *et al.* 2015; Chan *et al.* 2018; Li *et al.* 2019). Consequently, this has led to an increase in urban impermeable surface areas with severe road water, runoff pollution, and ecological damage (Paule-Mercado *et al.* 2017; Hou *et al.* 2019; Li *et al.* 2019). Meanwhile, the frequent urban flood inundation causes unavoidable disruption of transport and economic losses (Ruin *et al.* 2008; Lv *et al.* 2018). Nonetheless, monitoring the urban flood inundation can minimize the damages and losses (Versini 2012). Monitoring of the urban road inundation plays a key role in the application of urban flood inundation evaluation. Therefore, it is vital to monitor the urban road inundation to avert urban flood disaster.

Conventional urban flood inundation (e.g., road inundation) measurement methods (e.g., manual measurement, auxiliary mark method) have demonstrated several disadvantages under complicated climate and topographic surroundings, including insecurity, time-consuming, and high cost (Nair & Rao 2016; Zhang *et al.* 2019). In contrast with the traditional manual measurement mode, the sensors in modern measurement systems exhibit high precision. Nevertheless, the sensors might be damaged and buried by frequent flood events (Lin *et al.* 2018). Also, the measurement readings could be affected by local electricity supply and Internet access (Amin 2011). Additionally, deep learning techniques have been effectively used in object recognition. For instance, they are widely used for the automatic detection of objects in water (Kang *et al.* 2018; Cheng *et al.* 2019; Zhou *et al.* 2019). Notably, deep learning technology has a better accuracy advantage compared with conventional artificial neural networks; besides, it utilizes the available large amount of unlabeled data (Wu *et al.* 2015). The deep learning technique has demonstrated excellent capabilities to automatically learn complex and key features from raw data with better accuracy in image object detection (Yu *et al.* 2019; Zhou *et al.* 2019). Meanwhile, the advanced deep learning technique can effectively help to establish accurate prediction models (Le *et al.* 2020). Thus, unlike the conventional manual measurement and sensors in urban flood inundation (e.g., road inundation) monitoring, the deep learning technique extracts feature information of objects with low-cost, security, and satisfactory performance.

In recent years, the potential deep learning frameworks have applied Convolutional Neural Networks (CNNs; Srivastava *et al.* 2015) to train the network, including Faster R-CNN (Ren *et al.* 2017), Mask R-CNN (Yu *et al.* 2019), SSD (Sun *et al.* 2020), R-FCN (Si *et al.* 2019), and YOLO (Redmon & Farhadi 2017). These methods were well applied in image automatic recognition and classification problems (Van *et al.* 2020). Out of these, YOLO is a state-of-the-art framework for object detection and classification with a very deep layer and special residual net (Redmon *et al.* 2016; Zhang *et al.* 2020). The object detection can directly be evaluated by image pixels as a single regression problem using YOLO, the bounding boxes, and class probabilities. A single CNN in YOLO simultaneously

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

predicts multiple bounding boxes and their class probabilities (Koirala *et al.* 2019). CNN harbors an intelligent learning mechanism, hence easier classification or prediction of objects, and learning the essential features of the object images from a small number of samples (Geng *et al.* 2018). The original YOLO network was referred to as YOLOv1. Furthermore, it was enhanced to YOLOv2 based on the YOLOv1 model. Therefore, YOLO v2 is also a state-of-the-art detection framework, with improved detection speed under stable accuracy. The great speed and accuracy of YOLOv2 was an improvement of YOLOv1, it uses a pass-through layer, higher resolution classifier, and anchor boxes (Redmon *et al.* 2016). Scholars confirmed that YOLOv2 has an advantage in image object detection (Redmon & Farhadi 2017; Arcos-García *et al.* 2018; Zhang *et al.* 2020). Besides, the current YOLOv2 can be inferenced in real-time and is still robust for object detection tasks compared with other methods (Ye *et al.* 2020). Based on the above analysis, this study used YOLOv2 due to its excellent detection accuracy and speed.

Nonetheless, limited studies have conducted urban road inundation detection using the deep learning technique (Rokni *et al.* 2015; Lin *et al.* 2018). For example, Rokni *et al.* (2015) developed a novel method of integrating pixel-level image fusion and image classification techniques on the lake surface water change detection. Elsewhere, Lin *et al.* (2018) resolved the automatic water-level detection of the river channels using the computer vision technique. Zhang *et al.* (2019) utilized the NIR-imaging video camera to obtain water-level measurements for rivers. The above-mentioned studies suggest that water-level information of rivers and lakes can be obtained using other techniques; however, these methods have not yet been used to investigate the urban road inundation. Previous studies explored inundation areas (Lv *et al.* 2018; Bhola *et al.* 2019), but a few limitations have been reported. For instance, Lv *et al.* (2018) developed a raindrop photometric model (RPM) that extracted information from an inundation region; however, the area was not computed, and the stability of the camera and rain can easily affect the implementation of this method. Bhola *et al.* (2019) adopted both deep learning and edge detection techniques to forecast flood inundation, but they primarily focused on identifying the water surface depths of a small river rather than inundated road. So far, no study has systematically detected urban road inundation based on deep learning techniques. Thus, this study aims to provide a novel idea for experimental urban road inundation automatic monitoring using the YOLOv2 deep learning framework based on the collected images dataset. We believe that this could easily produce better performance on urban road inundation accurate detection considering the different scene images, including experimental rainwater collecting tanks inundation and urban road inundation with water.

METHODOLOGY

To automatically identify the urban road inundation, this work applied the YOLOv2-based Darknet-19 network to extract inundation areas. Moreover, the camera technology was applied to support the collection of images.

An automatic monitoring method for experimental urban road inundation based on the deep learning technique

This work proposed an automatic monitoring method applied to experimental urban road inundation based on the YOLOv2 deep learning detection framework. The method was used to evaluate the water areas in the object region. The YOLOv2 framework was coded in the Python programming language using the available Python standard library. YOLOv2 was adopted as the object detector frame and the model was implemented on TensorFlow. TensorFlow object detection API was also used to complete part of the experimental setup. The structure of the urban inundation detection framework based on YOLOv2 is plotted in Figure 1. The results of object bounding boxes and class probabilities were predicted by full image pixels. An image was split into finer pixels then classified and used to generate inundation statistics. The features of collected images were extracted using initial convolutional layers of the network, and the last convolutional layer predicted the output probabilities and coordinates. Also, to further compute urban road inundation area, the model used hand-picked anchor boxes to predict bounding boxes based on the offsets of these anchors at every location in a feature map (Arcos-García *et al.* 2018). The *k*-means clustering on the

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

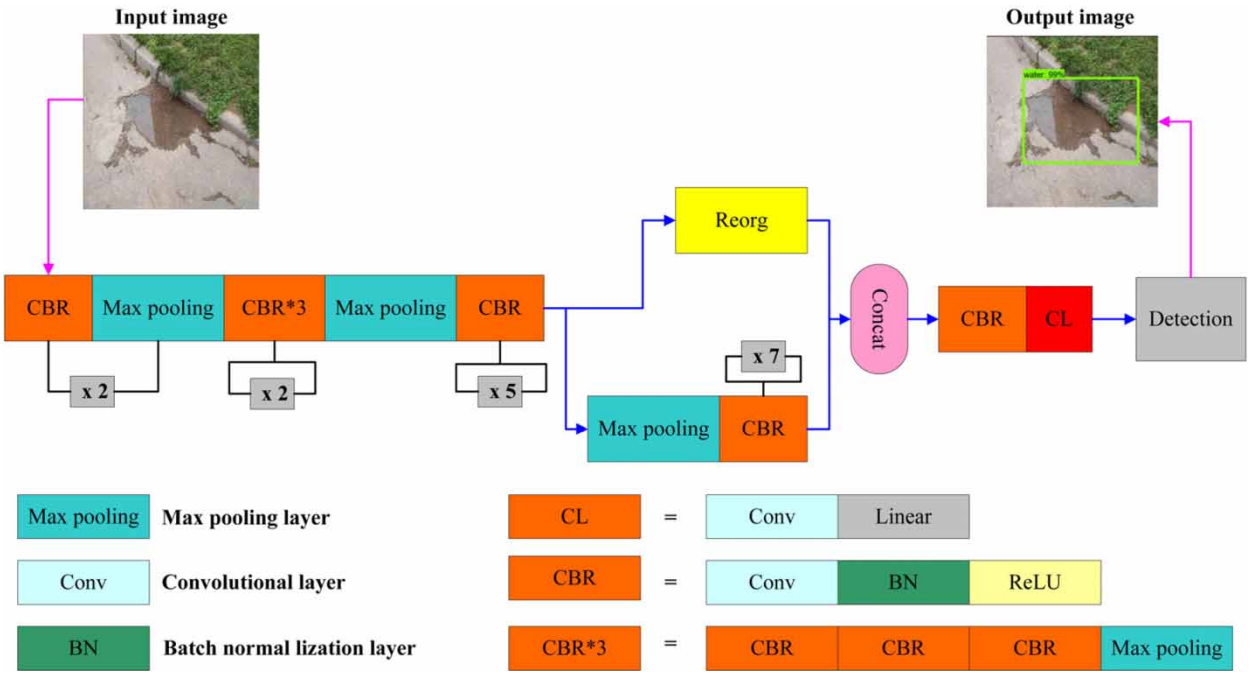


Figure 1 | The experimental urban road inundation detection framework is based on YOLOv2. 'Concat' refers to 'concatenation layer', 'Reorg' refers to 'reorganize', and 'Detection' refers to 'object detection'.

training set bounding boxes automatically predicted suitable anchor boxes (Redmon & Farhadi 2017).

The network of Darknet-19 is a novel classification model used as the base of YOLOv2 (Redmon & Farhadi 2017). Here, Darknet-19 was applied to perform object detection as a feature extractor. The Darknet-19 comprised 19 convolutional layers and 5 max-pooling layers, the fully connected layers were removed as shown in Figure 2. The characteristic of network structure was in alternating convolutional pooling layers with the organization of their neurons in a grid (Vasconcelos & Vasconcelos 2017). In the convolution layer, a few convolution kernels were equivalent to a set of linear filters used to obtain the features of the input image. After the convolution layer, pooling layers

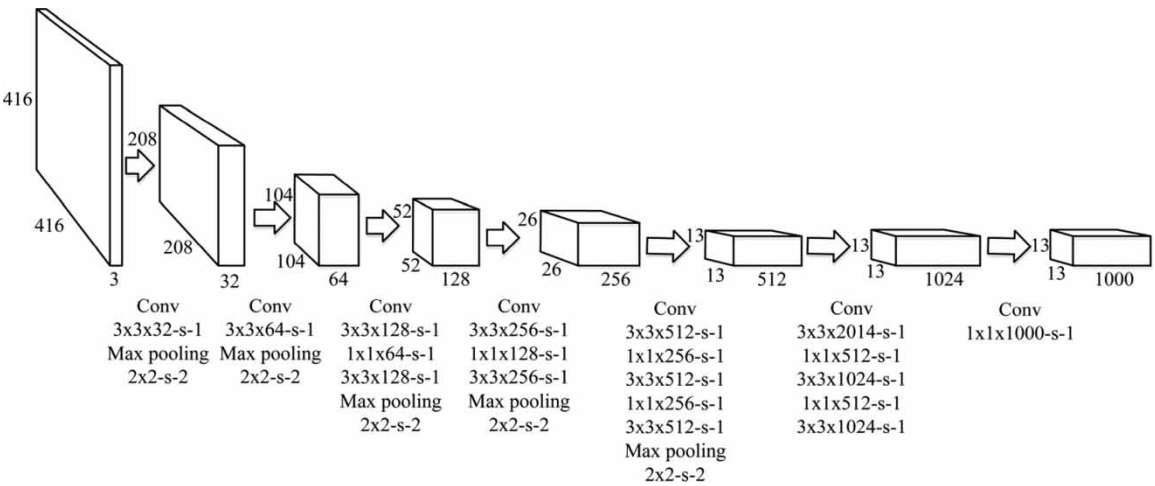


Figure 2 | Network structure of Darknet-19. 'Conv' refers to 'convolutional layer' and 'Max pooling' refers to 'Max-pooling layer'.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

down-sampled the output of the convolution layer, reducing a great deal of data processing while the important features image information was retained in this step (Geng *et al.* 2018). Thus, Darknet-19 doubled the number of feature maps after every pooling layer and mostly used 3×3 filters. Then, it adopted a global average pooling to predict with 1×1 filters, reducing the dimensionality of the feature space between 3×3 convolutions (Lin *et al.* 2014). Besides, the model used a batch normalization that regularized the model and improved model convergence, thereby providing better model computational performance and preventing overfitting in the model training process (Wang *et al.* 2018). Every batch normalization layer was applied after the convolutional layer and the last output of the activation function ReLU (Rectified Linear Unit) layer. Also, the nonlinear transformations function as ReLU was used to train the model, where the weights and variables of each layer were calculated during the training process.

As mentioned above, a comprehensive process of achieving experimental urban road inundation detection is shown in Figure 1. An original input resolution of 448×448 was used in the YOLO v2 model. To predict the objects with the addition of anchor boxes, the resolution was changed to 416×416 instead of 448×448 as reported by Redmon & Farhadi (2017). Therefore, this study also used an input resolution of 416×416. For the training of YOLOv2, the batch size was 32 while 5 was the type of anchor box size. The learning rate parameter was set as 0.001 while the epochs for training the network were 50. Notably, all the images trained together were collectively called one epoch. First, it was necessary to collect the experimental urban inundation images dataset before training the model on these images. Also, the object inundation region with water in the collected experimental urban inundation images was labeled before model training, and then these labeled images were used to train the model. Therefore, it was important to make the correct labels of the water object in the images, marking the location and labels for an object within the images, and reshaping the original images into 2D image format. The two consecutive layers of convolution and max-pooling had 3×3 convolutions (Figure 2). Finally, the network was modified for detection by removing the last convolutional layer and replaced by adding three 3×3 convolutional layers with 1,024 filters each followed by a final 1×1 convolutional layer with the number of outputs for the object detection (Arcos-García *et al.* 2018). The output results described the confidence scores and bounding boxes with green color for each input image. Overall, the automatic detection of experimental urban road inundation was established based on the YOLOv2, where the feature information of water position and the predicted anchor boxes in images were automatically identified.

Inundation area computation approach

As mentioned before, the proposed novel idea was used to obtain information for the experimental urban road inundation. The inundation area was computed for the images captured by the cameras in the vertical orientation. Additionally, if the images were collected from different angles, the object images were converted into an overlooked perspective as new test images using the inverse perspective transformation technique (Kim 2019), and then transformed images were used to estimate the inundation area. The inverse perspective transformation method transforms a two-dimensional image into a three-dimensional real-world space image, generating a bird-view (or top-view) image. The spatial calibration was performed using this image processing method to correct the barrel distortion of images; it is because some images are taken by a surveillance camera which is equipped with a wide-angle lens. Given that a different inundation region size appeared in the images, the proposed different assessment methods were used to calculate the inundation area under two scenarios. Figure 3 shows the flowchart of the inundation area computation process. Moreover, the CNN algorithm was applied to accurately predict the number of more detection anchor boxes within images (Koirala *et al.* 2019). For the first scenario, if the inundation region with water was an irregular area, the inundation area was obtained from the accumulation area of smaller predicted detection anchor boxes that covered the entire water region in the images. The formulation of the inundation area is as follows:

$$S_1 = \frac{S}{M} \times N \tag{1}$$

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

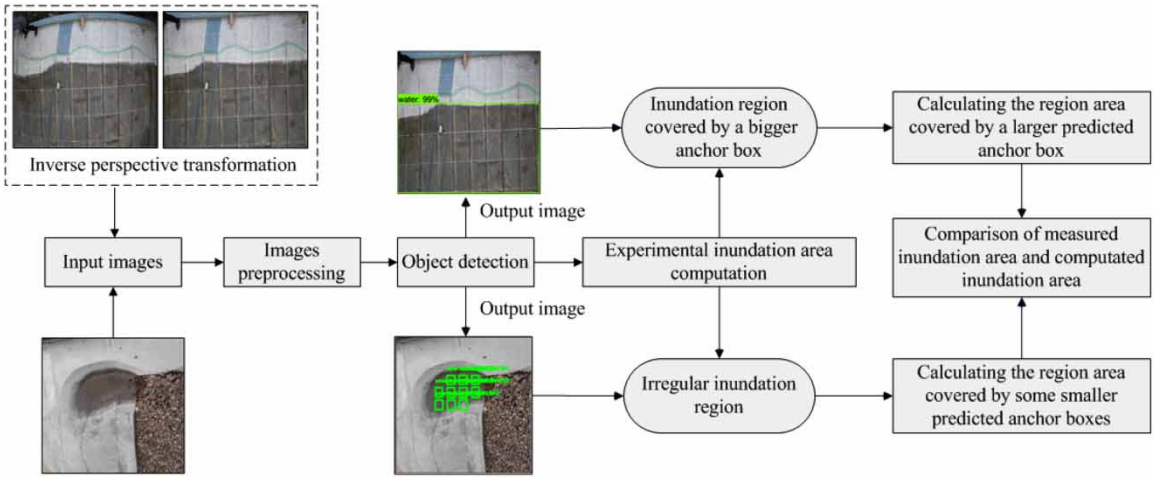


Figure 3 | The flowchart of the inundation area computation process.

where S_1 represents the computed inundation detection result area covered water in the region of each image; S represents the measured inundation area covered water in the region of each image reflecting the actual scene; N represents the number of predicted detection anchor boxes covered water; and M represents the number of predicted detection anchor boxes which covered the region of the entire image.

Furthermore, anchor box coordinate prediction was directly provided by the YOLOv2 detection framework based on labeled images dataset (Redmon & Farhadi 2017; Pi *et al.* 2020). After the model learned to predict objects in the images with an anchor box directly from the images, the coordinate values of the box covered water in each image were calculated. For the second scenario, if the inundation region with water was nearly covered by a predicted larger anchor box, the inundation area was evaluated by the area of this box coverage water region, reflecting the inundation region in the actual scene. Thus, the inundation area is also calculated using the formula:

$$S_2 = \frac{(x_{\max} - x_{\min})(y_{\max} - y_{\min})}{x_I} \times S \tag{2}$$

where S_2 depicts the computed inundation detection result area covered water in the region of each image; S represents the measured inundation area covered water in the region of each image reflecting actual scene; x_I depicts the pixel values of each input image with the size of 416×416; $(x_{\min}, y_{\min})$ and $(x_{\max}, y_{\max})$ are the pixel values of the upper-left and lower-right points of the predicted anchor box in each image, representing the coordinates value of the predicted anchor box. Additionally, the origin coordinate (0, 0) was at the upper-left corner of the image region.

To evaluate the performance of the method using the collected images dataset, we used the relative error of the measured area with the computed area. The smaller the relative error indicated a better performance by the proposed method.

Experimental images dataset acquisition and preprocessing

Experimental image acquisition

Considering the reasonableness of the images dataset and the safety of collecting images for this study, datasets were set up to improve the detection accuracy. The experimental images were acquired from the scene of experimental road water-logging and low-lying land flooded region in the Xi'an University of Technology, Xi'an, China. Meanwhile, some actual road inundation images were also used in this work. The automatic detection of experimental urban road inundation was carried out under both dry and wet conditions on roads. In order to effectively learn and extract the detailed features information of object water, some dry roads were tested in this work. The outside temperature was 30 °C. The experimental images were collected using

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

a smartphone camera with high resolution in different periods, including morning and afternoon under varying light intensity, different road types, and reflection effect in water. Small memory per image collected by a smartphone camera was contributed to image storage and preprocessing. Therefore, all collected images were stored in JPEG format with a high resolution. A total of 1,000 original images were captured under different angles and positions (Figures 4 and 5). Besides, to justify the usefulness of the proposed method, a few images were collected using a high-definition camera with the larger inundation areas. To memorize the detailed water features information based on YOLOv2, a simple geographical environment with a clear border around the water in the images, avoiding the impact of the complex terrain environment on important feature extraction. The experimental inundation images with water were considered as the important training dataset to enhance the accuracy rate of the actual inundation detection test. Moreover, to verify the performance of the method, the actual raining urban road inundation images were selected for the model test (Figure 6). The following section describes the image preprocessing.

Image preprocessing

To enhance the performance of the model training and show reliable detection results, the image preprocessing was performed before the model training. For example, the resolution of the original image was $2{,}352 \times 1{,}568$ pixels. The original image was down-scaled to 416×416 pixels. To prevent overfitting of the model due to the similarity of images and improving the reliability and diversity of images, the numbers of experimental images dataset were expanded via image preprocessing. The color



Figure 4 | Example of experimental water-logging road images.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user



Figure 5 | Example of experimental low-lying land flooded region images.



Figure 6 | Example of actual road inundation images.

transformation, image rotation, and salt and pepper noise removal methods were applied to make an expansion processing on images numbers from 1,000 to 3,000. Among them, image rotation and salt and pepper noise removal methods were applied to make an expansion processing on images numbers from 700 to 2,000. The images

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

numbered from 300 to 1,000 were expanded by the color transformation method. The acquired 3,000 images were subdivided into two groups, including model train dataset and validation dataset with the number of images being 2,500 and 500, respectively. The validation dataset comprised 100 actual raining urban road inundation. Besides, added 50 samples were used to evaluate the inundated area. Figure 7 shows the preprocessing for the original image to the preprocessed image. Besides, the training images first need to be marked, then the position and features information can be memorized by the model. The water region of the images as the detection object was labeled with the box using the labeling software. The labeled water images are shown at the bottom of Figure 7. Moreover, the model input requirement was primarily composed of down-scaled images dataset and XML files. These preprocessed images as the model input were applied in the model training. The information of coordinates and inundation region with water in the labeled image was described and saved as an XML file.

EXPERIMENTAL RESULTS AND DISCUSSION

The detection results of experimental urban road inundation based on YOLOv2 are described in the 'Experimental recognition accuracy evaluation' section; besides, we provided the recognition accuracy rate of experimental object detection evaluation. The new method for evaluating the experimental inundation area considering two scenarios is introduced in the 'Evaluation of the inundated area' section.

Experimental recognition accuracy evaluation

The accuracy rates in recognition of the model training and validation with an apparent change pattern are shown in Figure 8. As shown, the recognition accuracy rates increased at the beginning then reached a steady state for both the validation and train curves after 30 epochs. It was apparent that both model training and validation were higher than 0.9 and 0.8, respectively, and more than 10 for the model training epoch. As summarized in Table 1, the optimal training and

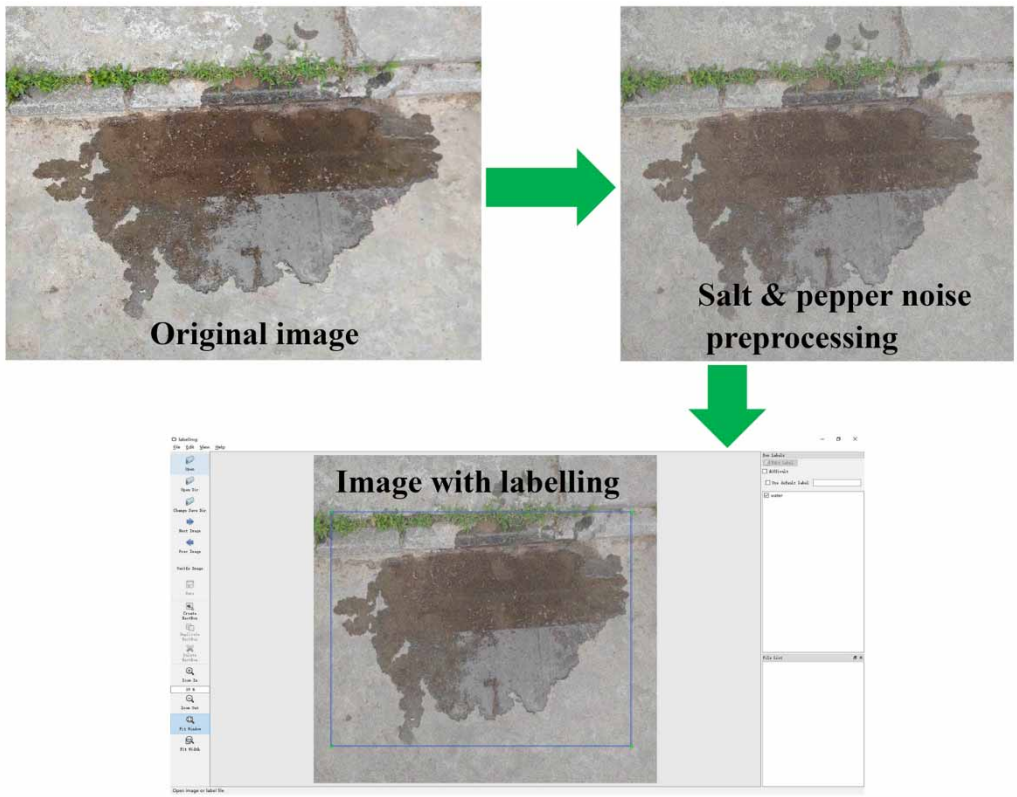


Figure 7 | Process of the images preprocessing.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

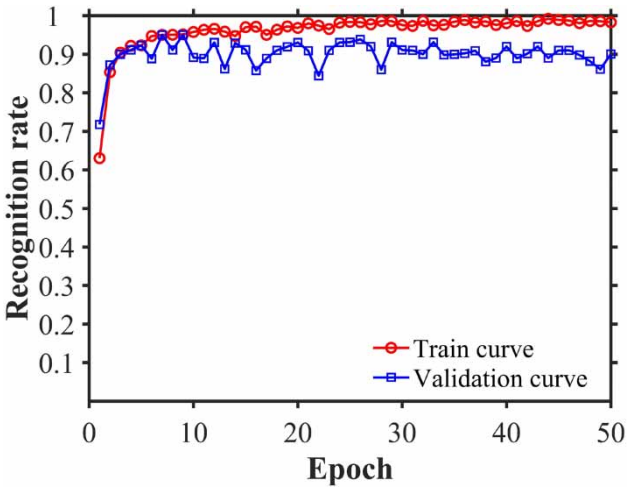


Figure 8 | Recognition rate curves of the model training and validation.

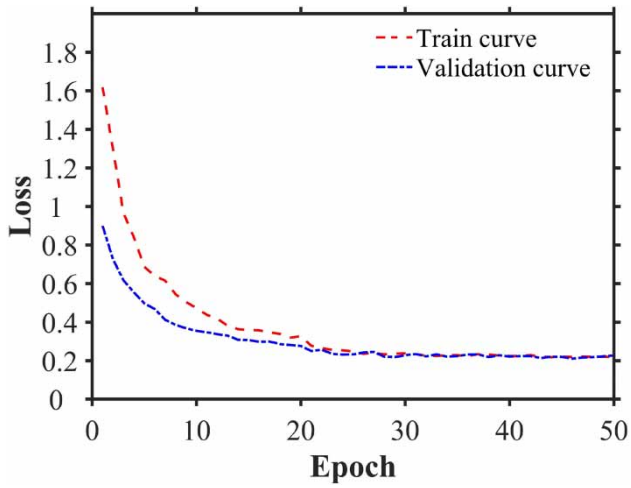


Figure 9 | Loss value curves of the model training and validation.

Table 1 | Accuracy rate evaluation

Network	Iterations number	Average accuracy rate for model training	Average accuracy rate for model validation
Darknet-19	50	96.1%	90.1%

validation average accuracy rates were 96.1 and 90.1%, respectively, indicating a better performance for model training and validation under the dataset. This phenomenon reflected that a steady state appeared in two curves (Figure 8) with higher accuracy rates, indicating that the training and validation reached a satisfactory convergence state. Meanwhile, there was a small error of approximately 10% compared with the model validation average accuracy rate (90.1%). Notably, weather factors (Lv *et al.* 2018) or human labeling errors in the image preprocessing potentially affected the validity of the algorithm (Koirala *et al.* 2019). Moreover, as shown in Figure 9, the loss values of the model training and validation gradually decreased at the beginning then reached a steady state. Again, the deviation in prediction loss of the model gradually decreased when the loss function of the small sample batches kept updating during the training process (Yu *et al.* 2019), hence a better performance of model training.

The detection findings of experimental road inundation, experimental low-lying land flooded region, and actual road inundation based on the model under a similar training images dataset are illustrated in Figures 10–12. The results revealed that the object inundation recognition exhibited a higher degree of confidence based on Figures 10–12. The method of water inundation detection performed better with a greater image recognition accuracy rate. Moreover, it showed the highest detection accuracy rate (99%) in the upper-left corner of the anchor box with green in these images (Figures 10 and 11). Generally, any classifier achieved an accuracy rate of 100% with difficultly, due to the effects of the lights, shadows, and complex obstacles (Geng *et al.* 2018). Meanwhile, the model recognition results were potentially influenced by human errors when labeling the training image dataset (Yu *et al.* 2019). Therefore, the findings suggested a satisfactory rationality in automatically extracting the inundation features using the proposed method and was examined and certified.

Figures 11 and 12 reveal the appearance of water accumulating in the low-lying region of the roads. The dimension of detection boxes with green changed automatically based on the size of the inundation region covered by water. The finer validation for the actual road inundation detection had a better accuracy rate (Figure 12). The model test images were from the actual road inundation scenes (Figure 6). The results showed that the established model for inundation detection could be also applied to automatically extract the appearance boundary features via model training and autonomous learning. Higher recognition results indicated that the method had a significant and effective detection performance for inundated urban roads. Rokni *et al.* (2015) reported a satisfactory performance of detecting

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user



Figure 10 | Experimental low-lying land flooded region detection results. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/hydro.2021.156.

the lake surface water change in the global change detection, hence, confirming that the result of surface water detection effectively managed flood monitoring and warning. However, smaller-scale surface water was not considered. Therefore, these findings guide the monitoring of urban road inundation based on our automatic detection analysis.

Evaluation of the inundated area

Scenario 1

A reliable result of the high performance of the established model for the detection of the inundated region is provided in the previous section. In this subsection, the novel proposed method was used for computing inundation areas covered by water via adjusting the size of predicted anchor boxes when the inundation region with water had an irregular area scenario. For example, Figure 13 shows the measured water area and recognition result with 10 anchor boxes. Besides, a green mark above these boxes shows the confidence score and classification information for model output detection result. As shown in Figure 13(a), the water shape approximated ellipse in the image with the dimension of 53 cm × 29 cm and the inundation area were calculated by the ellipse area formula approximately 1,207.1 cm^2.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user



Figure 11 | Experimental road inundation detection results. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/hydro.2021.156.

Figure 13(b) shows a scenario where the entire inundation region with water was almost covered by 10 boxes in the image. This was because the fixed small size of detection anchor boxes was defined before model training.

Based on Figure 13, the measured area of the real image region had a dimension of 115 cm × 115 cm, covered by 121 detection anchor boxes under this scenario. Therefore, the inundation area was obtained via the statistic number of detection anchor boxes, as shown in Equation (1). The area of a detection box was 109.3 cm^2, and the entire covered inundation area of 10 detection boxes was about 1,093 cm^2. A comparison was conducted between the computed detection area based on the model and the measured area of the image water region with the relative error being approximately 10%, indicating the feasibility of the method and a good reliability for inundation area evaluation. In contrast with the inundation area of the traditional measurement method (Lin *et al.* 2018), we confirmed that the proposed method quickly assessed the inundated region area with accuracy. Above all, the deep learning technique for inundated region detection obtained the information of an area and extracted precise object features via training and autonomous learning. Moreover, the bad weather potentially affected the model detection precision of the results due to the object surroundings reflection and bright light mixed in the image.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user



Figure 12 | Actual road inundation detection results. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/hydro.2021.156.

Scenario 2

For the second scenario, the inundation area was obtained to compute the predicted box region area covered by water when the inundation region was nearly covered by a predicted larger anchor box. Fifty cases were tested in this section, and the experimental 20 images with the larger inundated region area were collected using a surveillance camera with high resolution. For instance, the example of four typical cases is shown in Figure 14, there was a better accuracy rate in the inundation automatic detection. The images were captured under a similar experiment site and taken in the vertical orientation (Figure 14(a) and 14(b)). Also, after the object images were converted into an overlooked perspective, the larger inundated region area was accurately computed as shown in Equation (2) (Figure 14(c) and 14(d)). Therefore, the total measured area in the real scene for these images was directly measured with the image input pixel of 416×416. Moreover, the predicted values of $(x_{\min}, y_{\min})$ and $(x_{\max}, y_{\max})$ coordinate of the anchor box for two images are listed in Table 2, which calculated the area proportion of the detection box coverage area to the total measured area, with the origin coordinates (0, 0) being at the upper-left corner of the image region.

Thereafter, the inundation area in the images is computed by Equation (2). Unlike the measured and computed areas for the inundation region with water, the relative

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

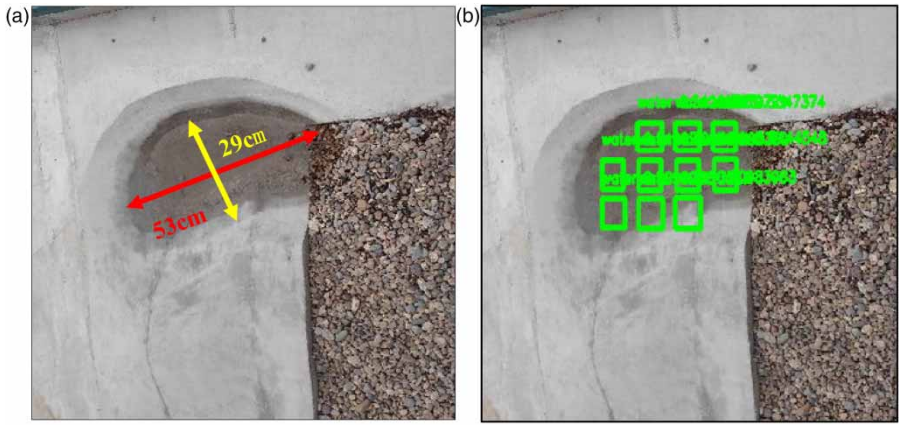


Figure 13 | Evaluation of the inundation area. (a) Measured area. (b) Recognition result with boxes. Please refer to the online version of this paper to see this figure in color: https://doi.org/10.2166/hydro.2021.156.

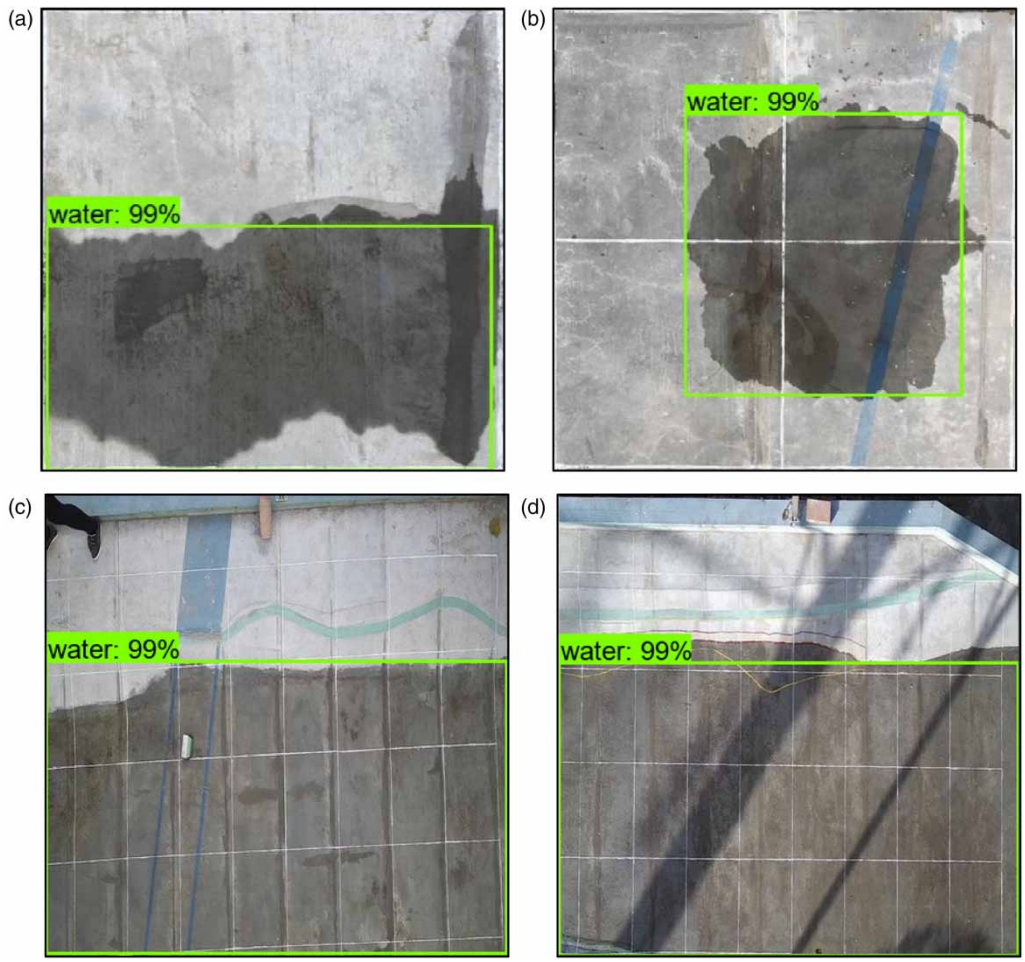


Figure 14 | Example of the inundated region area evaluation.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

Table 2 | Coordinate information of the anchor box and inundation area computation error

No.	x_{max}	x_{min}	y_{max}	y_{min}	Measured area (cm^2)	Computed area (cm^2)	Error percentage (%)
A	405.65	4.15	416.00	195.02	1,434	1,551	8.2%
B	366.19	118.02	348.55	94.17	3,850	4,414	14.6%
C	415.95	1.00	413.83	151.73	74,600	77,049	3.3%
D	415.27	0.83	416.00	151.77	83,400	77,581	7.0%

error percentages of 50 cases are illustrated in Figure 15, with the average relative error percentage of 7.7%. It is clear that the relative error percentages of all cases were less than 20%. In addition, the detailed coordinate information of the anchor box and the inundation area computation error of four typical cases are listed in Table 2, showing that the relative error percentage of image A, B, C, and D were 8.2, 14.6, 3.3, and 7.0%, respectively. The permissible smaller error showed a satisfactory prediction and area evaluation for the proposed method under the second scenario. Some of the explanations for the errors of inundation area evaluation could be affected by the shaking of the general camera (Lv *et al.* 2018). Meanwhile, the wet road or low-lying land flooded region with no inundation potentially affected the result detection, causing an error in evaluation of the inundation area. It is also possible that the varying light reflection of the water surface and complex geometry shape of water border could cause some errors. Above all, the experimental findings confirmed that the proposed idea could be applied to complete the inundation area computation with a better performance. Based on the above-mentioned analysis, our study produced a more complete evaluation for inundation area computation under two scenarios, hence providing an efficient strategy for monitoring urban road inundation.

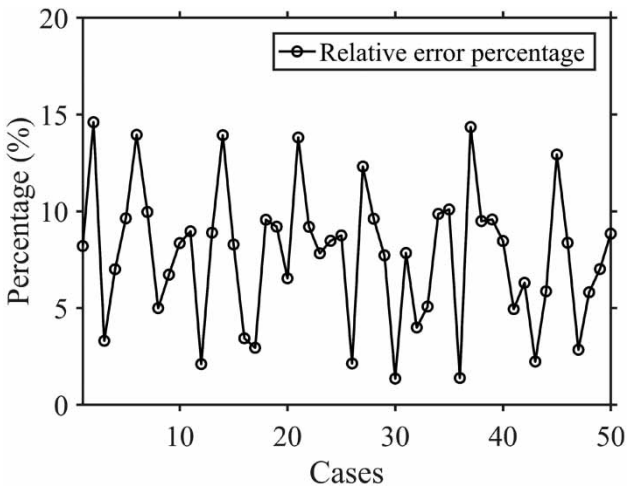


Figure 15 | Curve of relative error of the measured area with the computed area.

CONCLUSION

To facilitate the current urban road inundation automatic monitoring, a novel idea based on the YOLOv2 deep learning frame was applied to inundation region automatic detection and area computation. The complete process included image acquisition, preprocessing, inundation recognition, and the inundation area computation. Through analyzing the results, some conclusions are as follows:

- The proposed method based on the YOLOv2 deep learning framework could be effective for experimental flood inundation detection. The trained model exhibits better universality for different types of images dataset with varying angles and road conditions.
- Moreover, the results of inundation recognition accuracy rates showed that the model training and validation accuracy rates were high with 96.1 and 90.1%, respectively. Moreover, further validation confirmed that it harbors a higher accuracy rate for the actual road inundation detection. Therefore, the results indicated the impact of high accuracy and reliability for road inundation automatic detection.
- Furthermore, by comparing the measured inundation area and computation inundation area for two scenarios, the area relative errors of the test cases were less than 20%, with the average relative error percentage of 7.7%. The findings indicated an effective performance in the assessment of the inundation area.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

Summarily, our findings revealed a higher accuracy and an efficient feasibility of the proposed method in detecting experimental urban road inundation and computing the area. These results guide the urban flood-warning and road inundation monitoring. However, limitations existed during the process of images dataset collecting, for instance, the experimental inundation images were not collected at night, under continuous rainstorms, or other harsh weather conditions due to the potential effects of additional uncertain factors. Besides, the method was only applied in inundation area computation images captured in the vertical orientation without considering different angles. These complex conditions might change the recognition accuracy rate, hence warrants further investigation. The proposed method could provide a foundation and expanded guidance for experimental urban road inundation evaluation under different geographical environments. Furthermore, different geographical factors could be considered to improve the performance of the proposed methods. A follow-up study using a higher precision and great stability of camera on the urban road inundation detection considering the more comprehensive environment is essential. Moreover, water area and depths of the inundated region play a key role in the urban flood management and flood-warning; however, this work only considered inundation water area evaluation. Notably, deep learning and computer vision techniques are not entirely automated in the process of water depths perdition, thus requiring manual intervention (Bhola *et al.* 2019). As such, the proposed strategy cannot fully replace the numerical hydrologic and hydraulic models. Nonetheless, the proposed method can be applied to investigate the applicability of the numerical model. In addition to the above-mentioned assumption, further enhanced investigations on the real-time performance of urban flood inundation area and depths for predicting and monitoring are necessary.

ACKNOWLEDGEMENTS

This work is partly supported by the National Natural Science Foundation of China (52079106, 52009104); Water Conservancy Science and Technology Project of Shaanxi Province (Grant No. 2017slkj-14); Shaanxi International Science, Technology Foundation of China (Grant No. 2017KW-014); and the National Key Research and Development Program of China (2016YFC0402704). We are also thank the editor and the four anonymous reviewers whose insightful and constructive comments helped us to improve the quality of the paper. Hao Han and Jingming Hou contributed equally to this work.

CONFLICT OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

Amin, M. S. 2011 Smart grid: overview, issues and opportunities, advances and challenges in sensing, modeling, simulation, optimization and control. *Eur. J. Control* **17** (5–6), 547–567. https://doi.org/10.3166/ejc.17.547-567.

Arcos-García, Á., Álvarez-García, J. A. & Soria-Morillo, L. M. 2018 Evaluation of deep neural networks for traffic sign detection systems. *Neurocomputing* 332–344. https://doi.org/10.1016/j.neucom.2018.08.009.

Baek, S. S., Choi, D. H., Jung, J. W., Lee, H. J., Lee, H., Yoon, K. S. & Cho, K. H. 2015 Optimizing low impact development (LID) for stormwater runoff treatment in urban area, Korea: experimental and modeling approach. *Water Res.* **86**, 122–131. https://doi.org/10.1016/j.watres.2015.08.038.

Bhola, P. K., Nair, B. B., Leandro, J., Rao, S. N. & Disse, M. 2019 Flood inundation forecasts using validation data generated with the assistance of computer vision. *J. Hydroinform.* **21** (2), 240–256. https://doi.org/10.2166/hydro.2018.044.

Chan, F. K. S., Griffiths, J. A., Higgitt, D., Xu, S., Zhu, F., Tang, Y.-T., Xu, Y. & Thorne, C. R. 2018 "Sponge City" in China – a breakthrough of planning and flood risk management in the urban context. *Land Use Policy* **76**, 772–778. https://doi.org/10.1016/j.landusepol.2018.03.005.

Cheng, S., Zhang, S. & Zhang, D. 2019 Water quality monitoring method based on feedback self correcting dense connected

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

convolution network. *Neurocomputing* **349**, 301–313. https://doi.org/10.1016/j.neucom.2019.03.023.

Geng, L., Sun, J., Xiao, Z., Zhang, F. & Wu, J. 2018 Combining CNN and MRF for road detection. *Comput. Electr. Eng.* **70**, 895–903. https://doi.org/10.1016/j.compeleceng.2017.11.026.

Hou, J., Han, H., Qi, W., Guo, K., Li, Z. & Hinkelmann, R. 2019 Experimental investigation for impacts of rain storms and terrain slopes on low impact development effect in an idealized urban catchment. *J. Hydrol.* **579**, 124176. https://doi.org/10.1016/j.jhydrol.2019.124176.

Kang, J., Park, Y. J., Lee, J., Wang, S. H. & Eom, D. S. 2018 Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems. *IEEE Trans. Ind. Electron.* **65** (5), 4279–4289. https://doi.org/10.1109/TIE.2017.2764861.

Kim, J. B. 2019 Efficient vehicle detection and distance estimation based on aggregated channel features and inverse perspective mapping from a single camera. *Symmetry* **11** (10), 1205. https://doi.org/10.3390/sym11101205.

Koirala, A., Walsh, K. B., Wang, Z. & McCarthy, C. 2019 Deep learning – method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agr.* **162**, 219–234. https://doi.org/10.1016/j.compag.2019.04.017.

Le, H. V., Bui, Q. T., Bui, D. T., Tran, H. H. & Hoang, N. D. 2020 A hybrid intelligence system based on relevance vector machines and imperialist competitive optimization for modelling forest fire danger using GIS. *J. Environ. Inform.* **36** (1), 43–57. https://doi.org/10.3808/jei.201800404.

Li, Q., Wang, F., Yu, Y., Huang, Z., Li, M. & Guan, Y. 2019 Comprehensive performance evaluation of LID practices for the sponge city construction: a case study in Guangxi, China. *J. Environ. Manage.* **231**, 10–20. https://doi.org/10.1016/j.jenvman.2018.10.024.

Lin, M., Chen, Q. & Yan, S. 2014 Network in network. In ICLR.

Lin, Y.-T., Lin, Y.-C. & Han, J.-Y. 2018 Automatic water-level detection using single-camera images with varied poses. *Measurement* **127**, 167–174. https://doi.org/10.1016/j.measurement.2018.05.100.

Lv, Y., Gao, W., Yang, C. & Wang, N. 2018 Inundated areas extraction based on raindrop photometric model (RPM) in surveillance video. *Water* **10** (10), 1332. https://doi.org/10.3390/w10101332.

Nair, B. B. & Rao, S. 2016 Flood water depth estimation – a survey. In: *7th IEEE International Conference on Computational Intelligence & Computing Research*, Chennai, India. IEEE, pp. 1–4.

Paule-Mercado, M. A., Lee, B. Y., Memon, S. A., Umer, S. R. & Salim, I. 2017 Influence of land development on stormwater runoff from a mixed land use and land cover catchment. *Sci. Total Environ.* **599**, 2142–2155. https://doi.org/10.1016/j.scitotenv.2017.05.081.

Pi, Y., Nath, N. D. & Behzadan, A. H. 2020 Convolutional neural networks for object detection in aerial imagery for disaster response and recovery. *Adv. Eng. Inform.* **43**, 101009. https://doi.org/10.1016/j.aei.2019.101009.

Redmon, J. & Farhadi, A. 2017 YOLO9000: better, faster, stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9. https://doi.org/10.1109/CVPR.2017.690.

Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. 2016 You only look once: unified, real-time object detection. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788. https://doi.org/10.1109/CVPR.2016.91.

Ren, S., He, K., Girshick, R. & Sun, J. 2017 Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE T. Pattern Anal.* **39** (6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Rokni, K., Ahmad, A., Solaimani, K. & Hazini, S. 2015 A new approach for surface water change detection: integration of pixel level image fusion and image classification techniques. *Int. J. Appl. Earth Obs.* **34**, 226–234. https://doi.org/10.1016/j.jag.2014.08.014.

Ruin, I., Creutin, J.-D., Anquetin, A. & Lutoff, C. 2008 Human exposure to flash floods – relation between parameters and human vulnerability during a storm of September 2002 in Southern France. *J. Hydrol.* **361** (1–2), 199–213. https://doi.org/10.1016/j.jhydrol.2008.07.044.

Si, J., Lin, J., Jiang, F. & Shen, R. 2019 Hand-raising gesture detection in real classrooms using improved R-FCN. *Neurocomputing* **359**, 69–76. https://doi.org/10.1016/j.neucom.2019.05.031.

Srivastava, R. K., Greff, K. & Schmidhuber, J. 2015 Training very deep networks. *Adv. Neural Inf. Process Syst.* **28**, 2377–2385.

Sun, X., Gu, J. & Huang, R. 2020 A modified SSD method for electronic components fast recognition. *Optik* **205**, 163767. https://doi.org/10.1016/j.ijleo.2019.163767.

Van, S. P., Le, H. M., Thanh, D. V., Dang, T. D., Loc, H. H. & Anh, D. T. 2020 Deep learning convolutional neural network in rainfall-runoff modelling. *J. Hydroinform.* Published online. https://doi.org/10.2166/hydro.2020.095.

Vasconcelos, C. N. & Vasconcelos, B. N. 2017 Experiments using deep learning for dermoscopy image analysis. *Pattern Recogn. Lett.* 1–9. https://doi.org/10.1016/j.patrec.2017.11.005.

Versini, P.-A. 2012 Use of radar rainfall estimates and forecasts to prevent flash flood in real time by using a road inundation warning system. *J. Hydrol.* **416–417**, 157–170. https://doi.org/10.1016/j.jhydrol.2011.11.048.

Wang, J., Li, S., An, Z., Jiang, X., Qian, W. & Ji, S. 2018 Batch-normalized deep neural networks for achieving fast intelligent fault diagnosis of machines. *Neurocomputing* **329**, 53–65. https://doi.org/10.1016/j.neucom.2018.10.049.

Wu, Z., El-Maghraby, M. & Pathak, S. 2015 Applications of deep learning for smart water networks. *Procedia Eng.* **119**, 479–485. https://doi.org/10.1016/j.proeng.2015.08.870.

Ye, X., Hong, D., Chen, H., Hsiao, P. & Fu, L. 2020 A two-stage real-time YOLOv2-based road marking detector with lightweight spatial transformation-invariant classification. *Image Vision Comput.* **102**, 103978. https://doi.org/10.1016/j.imavis.2020.103978.

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user

Yu, Y., Zhang, K., Yang, L. & Zhang, D. 2019 Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. *Comput. Electron. Agr.* **163**, 104846. https://doi.org/10.1016/j.compag.2019.06.001.

Zhang, Z., Zhou, Y., Liu, H. & Gao, H. 2019 In-situ water level measurement using NIR-imaging video camera. *Flow Meas. Instrum.* **67**, 95–106. https://doi.org/10.1016/j.flowmeasinst.2019.04.004.

Zhang, J., Yang, X., Li, W., Zhang, S. & Jia, Y. 2020 Automatic detection of moisture damages in asphalt pavements from GPR data with deep CNN and IRS method. *Automat. Constr.* **113**, 103119. https://doi.org/10.1016/j.autcon.2020.103119.

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K. & Fu, G. 2019 Deep learning identifies accurate burst locations in water distribution networks. *Water Res.* **166**, 115058. https://doi.org/10.1016/j.watres.2019.115058.

First received 15 October 2020; accepted in revised form 30 April 2021. Available online 17 May 2021

Downloaded from http://iwaponline.com/jh/article-pdf/23/4/764/910488/jh0230764.pdf
by LOUGHBOROUGH UNIVERSITY user