Uncorrected Proof

Journal of Hydroinformatics

© 2021 The Authors

Journal of Hydroinformatics Vol 00 No 0, 1 doi: 10.2166/hydro.2021.118

Detecting, extracting and classifying foreign objects in inter-basin channels to ensure water supply safety

Junjie Chen[a,b] and Donghai Liu[a,*]

[a] State Key Laboratory of Hydraulic Engineering Simulation and Safety, Tianjin University, 135 Yaguan Road, Tianjin 300350, China
[b] Department of Real Estate and Construction, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, China
*Corresponding author. E-mail: liudh@tju.edu.cn

DL, 0000-0003-4509-2271

ABSTRACT

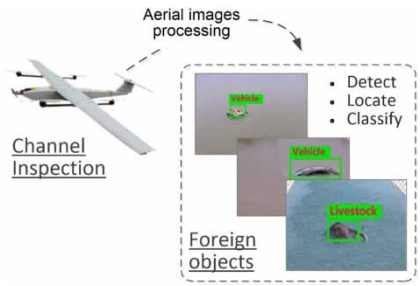
Foreign objects (e.g., livestock, rafting, and vehicles) intruded into inter-basin channels pose threats to water quality and water supply safety. Timely detection of the foreign objects and acquiring relevant information (e.g., quantities, geometry, and types) is a premise to enforce proactive measures to control potential loss. Large-scale water channels usually span a long distance and hence are difficult to be efficiently covered by manual inspection. Applying unmanned aerial vehicles for inspection can provide time-sensitive aerial images, from which intrusion incidents can be visually pinpointed. To automate the processing of such aerial images, this paper aims to propose a method based on computer vision to detect, extract, and classify foreign objects in water channels. The proposed approach includes four steps, i.e., aerial image preprocessing, abnormal region detection, instance extraction, and foreign object classification. Experiments demonstrate the efficacy of the approach, which can recognize three typical foreign objects (i.e., livestock, rafting, and vehicle) with a robust performance. The proposed approach can raise early awareness of intrusion incidents in water channels for water quality assurance.

Key words: computer vision, foreign objects, object detection, unmanned aerial vehicle (UAV), water quality, water supply safety

HIGHLIGHTS

- This study proposes an aerial image processing method to recognize foreign objects (FOs).
- To simultaneously detect, extract, and classify FO in water channel.
- To integrate superpixel segmentation and support vector machine for abnormal region detection.
- To propose a distance-aware algorithm to cluster abnormal regions.
- To develop a hierarchical voting mechanism for FO classification.

GRAPHICAL ABSTRACT



INTRODUCTION

Inter-basin water channels play an important role in allocating water resources across regions and maintaining the well-being of society (Zhang *et al.* 2015; Jia 2016). The operation of these water channels is threatened by occasionally occurring safety hazards (Norvanchig & Randhir 2021), such as water contamination, ice jam, and structural damages (Zhang *et al.* 2021).

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Although non-point sources such as stormwater runoff is notably an important risk factor (Wang *et al.* 2013; Norvanchig & Randhir 2021), the adverse influence of foreign object intrusion should not be ignored, especially for long-distance inter-basin water channels. Foreign objects are defined as those alien objects intruding into the water body. Typical foreign objects include livestock when the channel spans pastures (Lucey & Glennon 2021), vehicles fallen into the channel due to traffic accidents (Desh Deep 2019), and floating raft driven by external visitors (Farooquee *et al.* 2008). The intrusion of these foreign objects poses great threats to the channel operation. First, microbes carried by the livestock and petroleum leaked out from the vehicles can contaminate the water, thus degrading the water quality (Schumacher 2002). On the contrary, those external visitors breaking into the periphery of water channel for rafting can get drowned in the channel, raising subsequent judicial disputes.

To mitigate potential loss, timely detection of the intruded foreign objects is crucial. Current practice relies on inspectors or relevant personnel to patrol along the route to manually identify foreign objects. Such manual practice is labor-intensive and time-consuming. The deficiency is further enlarged by the fact that inter-basin channels usually span large geospatial areas (e.g., pasture, alpine region, or even desert) with enormous differences regarding natural and cultural characteristics. Recent advancements of unmanned aerial vehicles (UAVs) have potential to address the limitations. The high mobility of UAVs can significantly improve the efficiency by shortening the time spent on inspection and increasing the frequency of inspection. From aerial videos/photographs recorded by the airborne cameras, computer vision can be used to automatically detect foreign objects in a timely manner.

Equally important is to acquire relevant information of the intrusion incidents. For example, the quantity and geometry of the detected foreign objects can provide decision-makers with an overview of the scale and degree of severity of an incident, based on which a corresponding level of alarm can be issued. On the contrary, information about the types (e.g., livestock, vehicles, and rafting) of foreign objects is important for the determination of corresponding counter-measures, as different foreign objects bring different adverse impacts and maintain different characteristics. For example, when the detected objects are livestock bodies or vehicle fragments, they should be salvaged as soon as possible; however, if illegal rafting is identified, the break-in visitors should be ejected.

Detecting, extracting, and classifying objects of interest from images are not a unique problem in water quality management. Rather, the problem is widely encountered in various domains such as urban planning (Li & Yang 2020) and vegetation management (Carlier *et al.* 2020). In computer vision, it is usually formulated as a problem of object detection or instance segmentation, which aims not only to tell if certain objects exist, but also locate the pixel area of each object of different classes on given images. A solution to the problem is deep learning (DL), which has gained momentum since AlexNet's great success in the ILSVRC 2012 (Alom *et al.* 2018). However, the training of DL models relies on a massive amount of data, which are difficult or even impossible to collect. Annotating such magnitude of data is extremely laborious. For object detection, the annotation is done by designating bounding boxes around the objects of interest, which is already onerous enough. Let alone the time and efforts required to assign pixelwise labels in the scenario of instance segmentation. Even if the onerous data annotation jobs are doable, one might find it difficult to obtain sufficient images of the objects of interest from the outset, as they only appear when some infrequent accidental events happen. This is especially true in the scenario of water supply channels, where the occurrence of foreign object intrusion is an accidental event, and thus available aerial photographs of such adverse situations are usually too limited to allow the training of an actionable DL model.

To ensure water supply safety, this paper aims to provide a computer vision method that can automatically detect, extract, and classify foreign objects from aerial photographs. The proposed solution is expected to enable early awareness of the intrusion incidents and thus improve the capability of emergency response for water supply channels. The research is confined to the water channel inspection under normal operation. For extreme cases such as flooding, the special weather and hydraulic condition might pose additional challenges to aerial image collection and processing, which are not considered in this study. The contributions can be articulated from two aspects. From a theoretical perspective, this study makes contributions to the problem of vision-based foreign objects recognition from aerial images. It proposes a novel method including image preprocessing, abnormal detection, instance extraction, and foreign object classification to simultaneously detect, extract, and classify foreign objects under the constraints of data scarcity. The effectiveness of the method is demonstrated by comprehensive experimental studies. From a practical point of view, the proposed method provides a post-processing tool of aerial images to enable a more efficient practice of water safety management. With UAV and our method integrated, the timeliness and coverage of foreign objects detection can be significantly improved. The source code of our method has been shared on GitHub for possible reuse by the research community (civilServant-666 2021).

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

LITERATURE REVIEW

Timely detection of foreign objects in water body and acquiring their relevant information are important for ensuring water quality. The development of machine vision makes it possible to automatically detect foreign objects from images. Du *et al.* (2020) investigated the spatio-temporal pattern of total suspended matter in water by the combined application of time-series Landsat images and field survey. Lei (2019) compared the performance of three different object detection models, i.e., faster R-CNN, R-FCN, and SSD, in detecting float objects on man-made landscape lake. Lin (2020) integrated a series of image processing algorithms to detect and track floating foreign objects on urban river surface under complex background noises. Current studies mainly focused on detecting small floating contaminants (e.g., plastic bags, leaves, and float grass) on small water area (e.g., man-made landscape lake, and urban canal/river). However, such small contaminants generated from the municipal sector are not usually observed in the case of inter-basin water channels as these projects are normally located at remote regions. In contrast, more typical objects are those with bulky size such as livestock, fallen vehicles, and rafting, the detection of which has seldom been investigated. In addition, images processed by existing studies are usually collected by stationary or slow-moving platforms (e.g., autonomous patrolling ship) close to water surface. These methods are difficult to efficiently cover the range of large-scale water channels, as they usually span a long distance.

With the proliferation of commercial drones in recent years, the application of UAV has been gaining momentum. Because of its mobility, flexibility, and wide coverage, UAV is competent to undertake time-sensitive and hazardous tasks such as building and infrastructure inspection (Kim *et al.* 2018; Cai *et al.* 2019; Liu *et al.* 2019; Deng *et al.* 2020), power grid management (Jones 2007; Rengaraju *et al.* 2014; Bhola *et al.* 2018), progress monitoring (Wang *et al.* 2015), and construction safety evaluation (Melo *et al.* 2017). Computer vision has been used to automatically detect, extract, or segment objects of interest from aerial photographs collected by UAV. Kim *et al.* (2019) applied a deep neural network, YOLO-V3, to detect and localize construction resources at the jobsite for the prevention against struck-by hazards. Pi *et al.* (2020) investigated the application of CNN for the detection of ground objects such as roofs, cars, vegetation, debris, and flooded areas from aerial imagery. Zhu *et al.* (2020) presented a large-scale vehicle detection and counting benchmark dataset, which includes 15,532 pairs of RGB and infrared images collected by camera-equipped drones. Yu *et al.* (2016) proposed an approach to extract airplanes from satellite images based on the deep Hough-Forests method. Similar to the above scenarios, the integration of UAV and vision-based object detection can be expected to address the limitations of manual inspection in water supply channels. By exploiting the UAV-collected photographs with machine vision, foreign objects and their relevant information regarding quantities and geometry can be automatically identified in a timely manner.

DL requires large amounts of data to enable model training in an end-to-end manner. Such big dataset is difficult to collect in many scenarios including water supply channels. Under the constraints of limited available data, classical machine learning and image processing approaches can serve as potential alternatives. For example, Bhola *et al.* (2018) proposed a method integrating spectral clustering and spatial segmentation to detect power lines, which was evaluated on a small collection of aerial images. Carlier *et al.* (2020) investigated the feasibility of using morphological spatial pattern analysis to replace manual visual estimation for plant cover measurement on a collection of 30 images. Chen & Liu (2021) trained a bottom-up model for slope damage detection with a dataset consisting of <100 photos. Enlightened by these researches, this study intends to devise a foreign object detection approach suitable for small dataset by integrating image processing operations, feature handcrafting, and classical machine learning algorithms.

METHODOLOGY

Procedure of the proposed method

Figure 1 shows an overview of the proposed approach. To avoid the influence of irrelevant background objects (e.g., sky, vegetations, and surrounding buildings), regions of interest (ROI), i.e., water body, have been extracted from the input aerial images in advance. The process can be automated by the approach proposed in Chen *et al.* (2021), which aligns the geo-referenced aerial photos with a readily available 3D model for ROI extraction.

The proposed method includes the following four steps. First, preprocessing is performed to segment the input aerial images into superpixels, which converts the original images with macrofeatures into small subpixels with microfeature and expands the available training data. The second step addresses the question of whether foreign object intrusion has occurred. A support vector machine (SVM) is trained to iterate over all superpixels on an image for normal/abnormal binary classification. The third step addresses the question of how many foreign objects exist in an image. It does so by

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

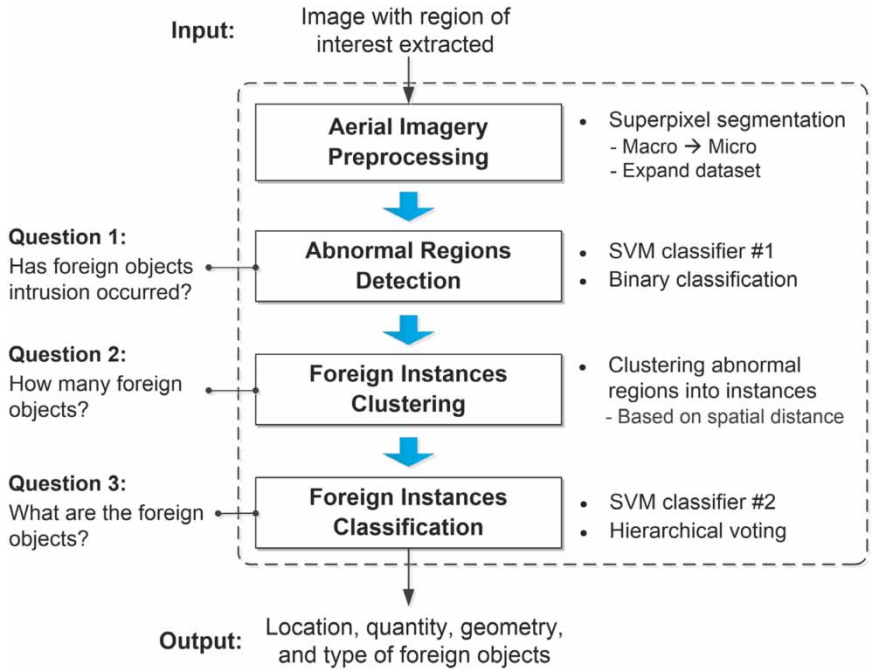


Figure 1 | Overview of the proposed approach.

clustering abnormal regions detected in the second step into individual instances based on their spatial distance. The fourth step addresses the question of what the detected foreign objects are. This is realized by training another SVM and devising a hierarchical voting mechanism. Following the above steps, the approach finally outputs information on whether intrusion incidents have occurred, and if so, how many foreign objects have been detected, where they are on the images, and the specific types of them.

Image preprocessing

The collected aerial photographs are first segmented into superpixels, which refer to small polygonal regions comprising neighboring pixels with similar color and texture (Ren & Malik 2003). The operation allows the extraction of microfeatures concerning either foreign objects or pure water. In addition, since subsequent model training is based on these small superpixels, it significantly expands the available amount of training data as compared to the original number of images before segmentation (Chen & Liu 2021). We use an algorithm called Simple Linear Iterative Clustering (SLIC; Achanta *et al.* 2012) for superpixel segmentation, because of its simplicity and robust performance. The algorithm includes the following four steps:

- Seed initialization: Evenly place initial seed points on the image based on a fixed interval.
- Seed points optimization: Re-assign pixels with the lowest gradients in the neighboring regions of the initial seed points as the new seed points.
- K-means clustering: Using the seed points after optimization in step (2) as initial clustering centers, perform K-means clustering to all image pixels.
- Improve the connectivity: Re-assign superpixels with small area to their neighboring superpixels by transversal from the upper-left corner of the image to its bottom-right corner.

Detecting abnormal regions by SVM

After superpixel segmentation, abnormal regions that contain foreign objects are detected from the aerial images. To do so, a SVM classifier (denoted by SVM classifier #1) is trained to determine whether individual superpixels on the input image contain foreign objects. SVM is a widely used machine learning model, the training of which aims to search for a hyperplane that

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

can best separate samples of different classes in the dataset (Yan *et al.* 2017). The model can produce precise prediction results based on relatively small datasets and therefore is adopted in this study for foreign object detection.

The SVM classifier #1 takes features extracted from a superpixel as input, and output a label indicating whether the superpixel is normal (water body) or abnormal (foreign objects). The extracted features comprise two parts, i.e., local binary pattern (LBP) and boundary indicator (BI). The LBP, proposed by Ojala *et al.* (2002), is an image descriptor for characterizing texture features. Since water body usually presents a simple but repetitive pattern of ripples, which is quite different from the texture displayed by foreign objects such as vehicles and rafting, the LBP can be used to effectively characterize such differences (Pietikäinen *et al.* 2000). Other than the LBP, an index called BI is also integrated as part of the feature. The BI is a Boolean value indicating whether a concerned superpixel is at the boundary of water channel. Such information on spatial layout is considered because the water body at the channel boundary is subject to the influence of reflection of near-shore objects, and thus, the corresponding superpixels might display a different texture pattern than that in the normal situation.

The features extracted from all superpixels in the image collection, together with their class labels, form a training set on which the SVM can be trained. After training, the model can be deployed in the testing/operational stage: given a newly collected image, it iterates over all the superpixels on the image for normal/abnormal classification, and then output a binary map with the detected abnormal regions highlighted.

Clustering abnormal regions into foreign instances

The binary map obtained from the last step tells image regions containing foreign objects, but cannot distinguish between different individual foreign objects (or the so-called instances). Two disconnected regions on the binary map can either correspond to two individual instances, or belong to an identical instance.

It is observed that disconnected regions of the same instance are usually closer from each other than those of different instances. Based on the observation, an algorithm is developed to cluster disconnected regions into individual instances according to their relative spatial distribution. The algorithm includes two sub-algorithms, i.e., adjacency matrix generation and instance annotation. Figure 2 shows the general procedure of the algorithm with an example.

Sub-algorithm #1: adjacency matrix generation

The spatial relationship between regions plays an important role in determining whether they belong to the same instance. Such a relationship can be formulated as whether the regions are close enough that they can be deemed 'adjacent'. From this perspective, sub-algorithm #1 intends to generate a matrix that reflects the adjacency relation between any two arbitrary regions on an image. More details of the algorithm can be found in Section S1 of the Supplementary Material.

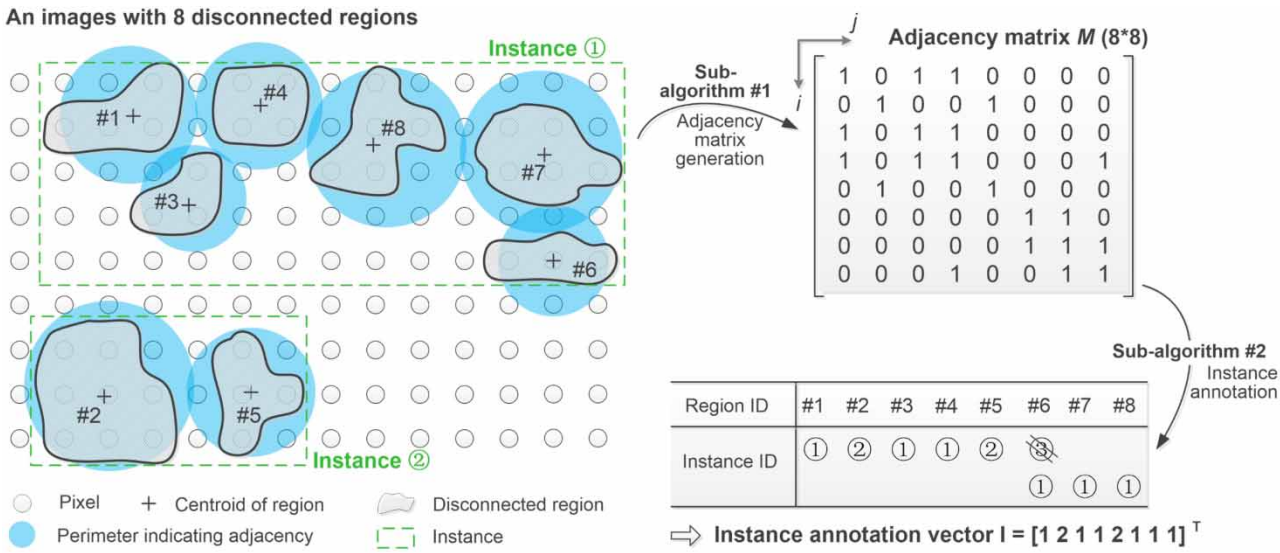


Figure 2 | Schematic diagram illustrating the general procedure of instance extraction.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Sub-algorithm #2: instance annotation

If two disconnected regions are adjacent to each other, they belong to the same instance. The adjacency relationship can be retrieved from the adjacency matrix M: when two rows in M have elements equal to 1 at the same column, the regions represented by the two rows are either directly adjacent to each other, or adjacent to an identical third-party region (see the example in Figure 2); thus, they belong to the same instance. Based on the above pattern, an instance annotation algorithm is devised to assign instance ID to each disconnected region for instance extraction. Refer to Section S2 of the Supplementary Material for more details of the algorithm.

Classifying foreign instances by a hierarchical voting mechanism

After extracting individual instances, the classes of these instances are recognized to provide actionable information for timely implementation of mitigation measures. According to project managers and stakeholders we have approached, small contaminants generated from the municipal sector are not usually observed in inter-basin water channels. In contrast, more typical objects are those with bulky size such as livestock, fallen vehicles, and rafting, which, therefore, will be considered as three candidates to be classified in this study.

Another SVM classifier (denoted by SVM classifier #2) is created, of which the input is LBP descriptors of an image patch occupied by a single or multiple superpixels, and the output is the type of foreign objects corresponding to the image patch. The training of SVM classifier #2 follows the strategy mentioned in Section S3 of the Supplementary Material.

A hierarchical voting mechanism is devised to decide the class of an instance based on the results given by SVM classifier #2, as shown in Figure 3. Suppose an extracted instance comprises N_{supix} superpixels. With SVM classifier #2, each of the N_{supix} superpixels can be classified as a specific type of foreign objects. For example, in the bottom layer of Figure 3, the four superpixels are, respectively, recognized as 'vehicle', 'vehicle', 'livestock', and 'rafting'. Let the superpixels be combined in pairs, which will then result in $C^2_{N_{\text{supix}}}$ combinations. Image areas inside the bounding boxes surrounding the superpixel combinations are successively fed to the SVM classifier #2 to obtain the corresponding classification results (as shown by the second layer from bottom-up in Figure 3). Likewise, one can select any r superpixels in the instance to form new combinations, and then perform classification. For an instance with N_{supix} superpixels, we can obtain $\sum_{r=1}^{N_{\text{supix}}} C^r_{N_{\text{supix}}}$ classification results for all possible superpixel combinations. Calculate the number of votes for each class, and the one with the largest number of votes is assigned as the class of the instance. Note that when the largest votes are observed at multiple classes, the output class is set as 'uncertain'. The hierarchical mechanism allows a comprehensive consideration of the instance features at different levels of details (from micro to macro), which can avoid bias brought by only taking the microfeatures of a single superpixel into account.

EXPERIMENTAL STUDIES

Data collection and preprocessing

A dataset was obtained to evaluate the performance of the proposed approach, which was mainly collected in two ways. First, we reached out to competent departments of two water supply projects in China, both of which have experience in using UAVs for channel inspection (Liu *et al.* 2019; Chen & Liu 2021), to obtain relevant photos they previously collected.

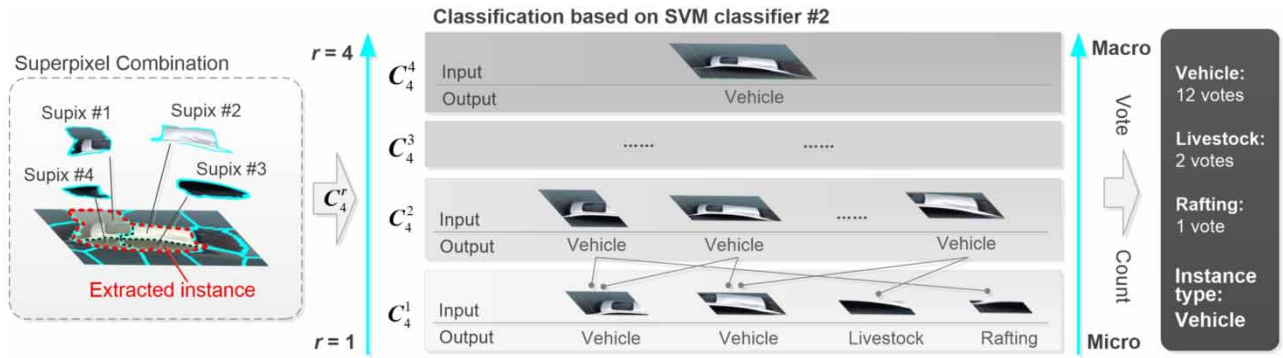


Figure 3 | Hierarchical voting mechanism for instance classification.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Second, online searching was performed on the Google, Baidu, and aerial photography forums. The dataset comprises 145 aerial images of water channels containing solely water body, and both water body and foreign objects. Section S4 of the Supplementary Material shows the statistical distribution of the dataset over different aspects of variations, demonstrating that the dataset has encompassed typical variations to ensure the model's generalizability.

The dataset was randomly split into a training set and a test set, respectively, for subsequent model training and evaluation. A 5-fold cross-validation is used for hyperparameter tuning on the training set. A training/testing split ratio of 8.5:1.5 was adopted by referring to previous studies (Cha *et al.* 2017; Chen & Liu 2021). Table 1 shows the detailed composition of the dataset. The collected images after background removal were segmented into superpixels with the SLIC algorithm. Data annotation was conducted to assign each superpixel to one of the following class labels, i.e., water, livestock, rafting, and vehicle, after which a dataset of superpixels can be obtained.

Results of abnormal region detection

The SVM classifier #1 for abnormal detection was trained on the 13,212 superpixels. The training was implemented with a Python machine learning library called scikit-learn. The quantity of 'water' superpixels (i.e., 12,716) far outnumbers that of the abnormal superpixels with foreign objects (i.e., 496). Machine learning models trained on such imbalanced dataset would tend to overfit the majority class, but perform poorly on the minority class. To mitigate such adverse effects, a higher penalty coefficient can be assigned to the minority class when training the model. With the scikit-learn library, this can be realized by setting the 'class-weight' parameter as 'balanced'. When training SVM classifiers, three hyperparameters need to be specified, i.e., kernel, C, and gamma. Values of the hyperparameters were determined by a 5-fold cross-validation, as introduced in Section S5 of the Supplementary Material. The performance of the calibrated SVM is compared with that of two other popular machine learning models, i.e., k-nearest neighbors (kNN) and decision tree (DT). As shown in Table 2, while the evaluation metrics for the 'normal' class are all at a high level for the three models, SVM is found to have attained the best 'abnormal' detection performance.

With the classification results given by the SVM model, abnormal regions on images from the test set can be detected, as shown in Figure 4. In each image pair, the first row is the original image with ground-truth foreign objects highlighted by green polygons, and the second row is the resulting binary image with abnormal regions detected. The model successfully detected the existence of foreign objects in all the abnormal image samples in the test set (Figure 4(a)). As for the seven normal image samples (Figure 4(b)), the model only misclassified a single image (#N-3) as containing foreign objects.

Table 1 | Detailed composition of the image dataset

Dataset split	Normal	Foreign objects intrusion			Total
	Water body	Livestock + Water body	Rafting + Water body	Vehicle + Water body	
Training	33	32	19	37	121
Test	7	6	5	6	24
Total	40	38	24	43	145

[a]The numbers in the table represent the quantities of images.

Table 2 | Abnormal detection performance of different machine learning models[a]

Metrics	SVM		kNN		DT	
	Abnormal	Normal	Abnormal	Normal	Abnormal	Normal
Precision	0.49	0.99	0.43	0.98	0.50	0.99
Recall	0.73	0.98	0.27	0.99	0.30	0.99
F1-score	0.58	0.99	0.33	0.99	0.37	0.99

[a]Performance of different models in the table has been calibrated by cross-validation.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

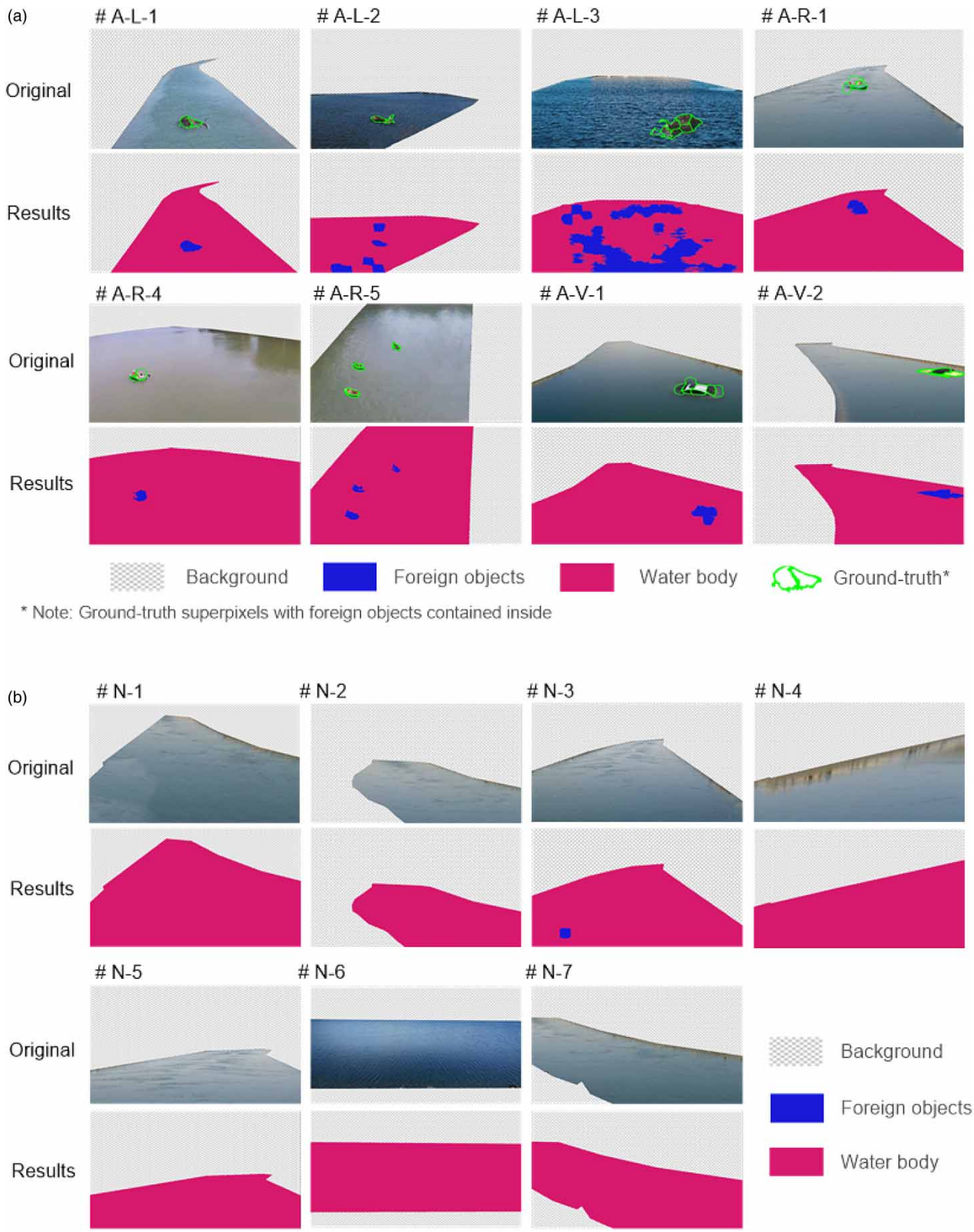


Figure 4 | Detected abnormal regions by SVM classifier #1 from (a) abnormal samples and (b) normal samples in the test set. Please refer to the online version of this paper to see this figure in colour: https://doi.org/10.2166/hydro.2021.118.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Results of instance clustering

The detected abnormal regions need to be clustered into individual instances of foreign objects. Figure 5 demonstrates results of instance extraction by applying the clustering algorithm proposed in the section 'Clustering abnormal regions into foreign instances'. The f in the figure is the scale factor in Equation (A2) in the Supplementary Material. With the increase of f, the algorithm tends to merge different regions into a handful of instances with large coverage (e.g., the most extreme case when $f = 2.5$ in Figure 5). However, if f is too small, the algorithm might fail to merge regions that are supposed to belong to identical instances, and even extract each region as an individual instance, as is observed when $f = 0.5$. When f is specified as 1.0, the algorithm yields the most reasonable results, as the two rafts in #A-R-3 were successfully separated and the two regions of an identical vehicle in #A-V-5 were correctly merged. Hence, the instances extracted with $f = 1.0$ will be used to evaluate the performance of foreign object classification in the subsequent section.

Results of instance classification

To classify the detected instance, SVM classifier #2 for foreign object classification was first trained. There are 496 individual superpixels of foreign objects in the training set. After combining superpixels belonging to identical instances, another 2,047 image patches were generated, thus expanding the original training set to one with $496 + 2{,}047 = 2{,}543$ samples. Similar to SVM classifier #1, cross-validation was performed to tune the hyperparameters of SVM classifier #2, as introduced in Section S6 of the Supplementary Material. After hyperparameter tuning, the selected optimal model was used to predict the classes of foreign instances in test images with the proposed hierarchical voting mechanism. As demonstrated by Figure 6(a), our approach successfully identified the classes of 80% (8 out of 10) of the foreign instances. For comparison, we also calculated the results without applying the hierarchical mechanism (neither train SVM classifier #2 with superpixel combinations nor vote for class hierarchically), which is listed in Figure 6(b). The accuracy is observed to drop significantly if the hierarchical mechanism was not applied.

Figure 6 only considers if an instance sample has been correctly classified. However, it is a common practice in the area of object detection and instance segmentation to also consider if a sample has been localized properly. To evaluate the localization precision, an index called intersection over union (IoU) is adopted, which is defined as follows:

$$IoU = \frac{R_{gt} \cap R_{pred}}{R_{gt} \cup R_{pred}} \tag{1}$$

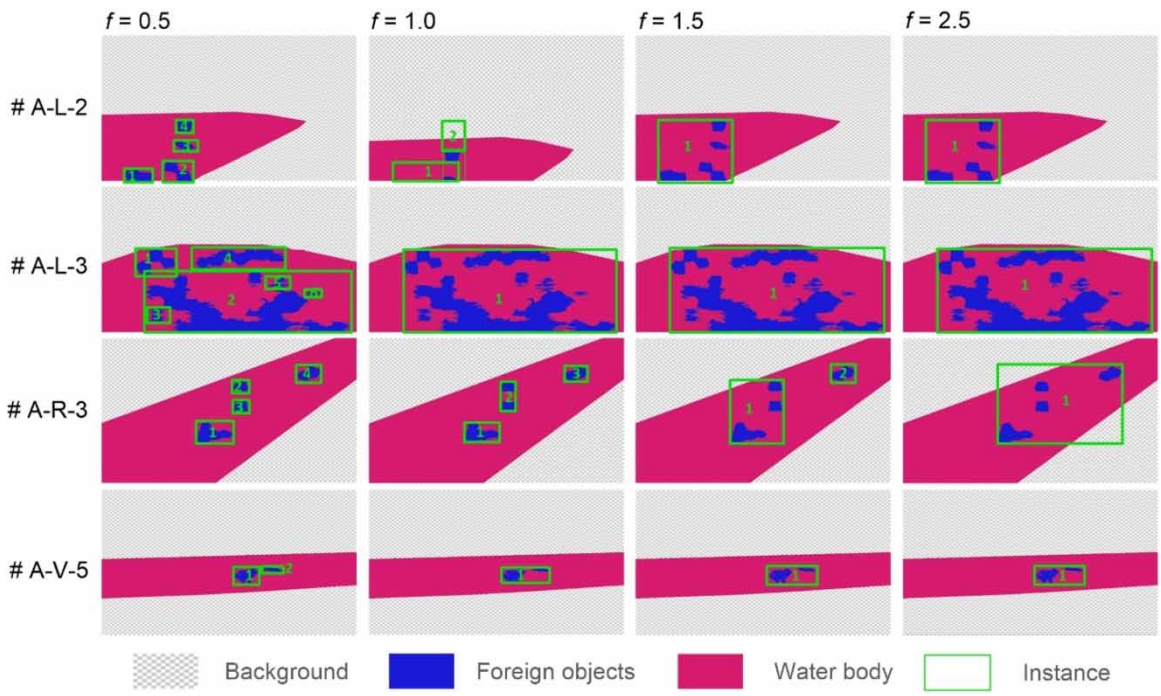


Figure 5 | Results of instance clustering under different scale ratios f.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

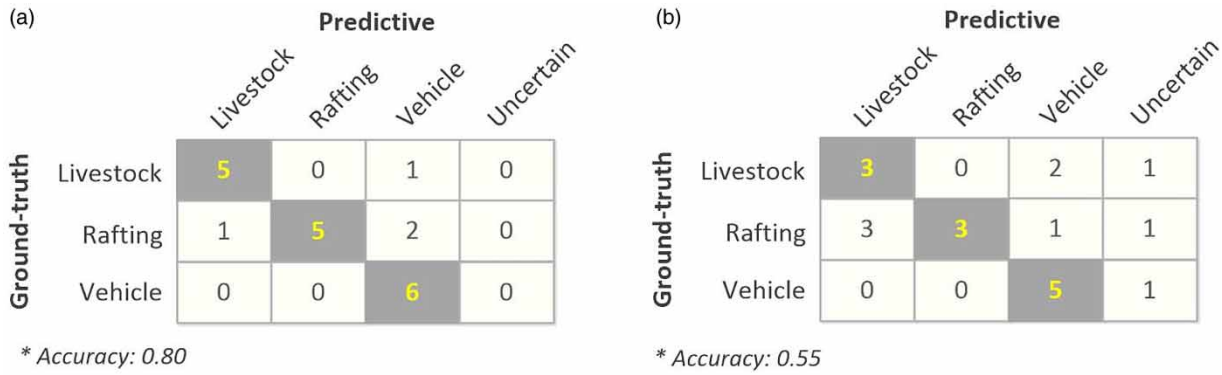


Figure 6 | Instance classification results (a) with applying the hierarchical voting mechanism and (b) without applying the hierarchical mechanism.

where R_{gt} and R_{pred} are respectively the ground-truth and predicted regions of foreign objects. In object detection, the region R is represented by bounding boxes. In instance segmentation, it is an irregular polygon occupied by corresponding labeled pixels. As widely accepted by previous studies (Rezazadeh Azar & McCabe 2012; Pi *et al.* 2020), an IoU value not <50% indicates a correct localization. By combining the classification metric and localization metric together, an instance of foreign objects is deemed as a true positive (TP) only when it is classified correctly, and reaches an IoU of not <50%; otherwise, it would be considered as a false negative (FN). A detection is deemed as a false positive (FP) when it is originally water body but incorrectly identified as the abnormal region.

Figure 7 shows predicted bounding boxes and classes of the detected instances (including those incorrectly detected) on the test images. Section S7 of the Supplementary Material provides detailed performance metrics of all the foreign instances in the test samples. Recall, or also called true positive rate in some cases (Azar & McCabe 2012; Rezazadeh Azar & McCabe 2012), is defined as 'TP/(TP + FN)', whereas precision is defined as 'TP/(TP + FP)'. Our approach achieved 70% for both recall and precision, while the false positives per image is 0.25 (=6/24). The proposed approach goes further in the level of granularity by providing boundary information of the detected instances. Figure 8 shows pixelwise results of instance segmentation, where we used all the detected superpixels in an instance as its geometry and considered class of the instance as its semantic label.

Time performance

The efficiency of the proposed approach at each step is evaluated in this section. The most two time-consuming tasks are, respectively, SLIC segmentation in the preprocessing stage and LBP calculation in the detection stage. The former on an average takes 2.7 s to segment an image into superpixels, while the latter costs 30 s per image, given it comprises 300 superpixels, for LBP feature calculation. The training of SVM classifier #1 and classifier #2 is quite efficient, which took <1 s (0.84 + 0.12 = 0.96 s) in total. An important factor to consider when evaluating the efficacy of an algorithm is its consuming time to process each image in the deployment stage. Given a newly collected image comprising 300 superpixels and containing a foreign instance made up of four superpixels, it will first consume 2.7 s at the preprocessing stage, 30 + 0.03 = 30.03 s for abnormal detection, 0.0003 s for instance extraction, and 1.5 + 0.06 = 1.56 s for foreign object classification, which sums up to 34.29 s in total.

DISCUSSION

Computer vision-enabled foreign objects detection

The study contributes to the problem of vision-based foreign object detection from the following aspects. On the one hand, the adopted superpixel segmentation technique breaks down the original problem of processing the entire image into hundreds of small problems of superpixel classification. This reduces the complexity of the problem and expands the original dataset by hundreds of times. On the other hand, the involvement of human knowledge mitigates the reliance on data volume. A high-level feature descriptor LBP was handcrafted to characterize the differentiation between water body and foreign objects, and among different types of foreign objects. In the instance extraction stage, an unsupervised clustering algorithm was provided

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

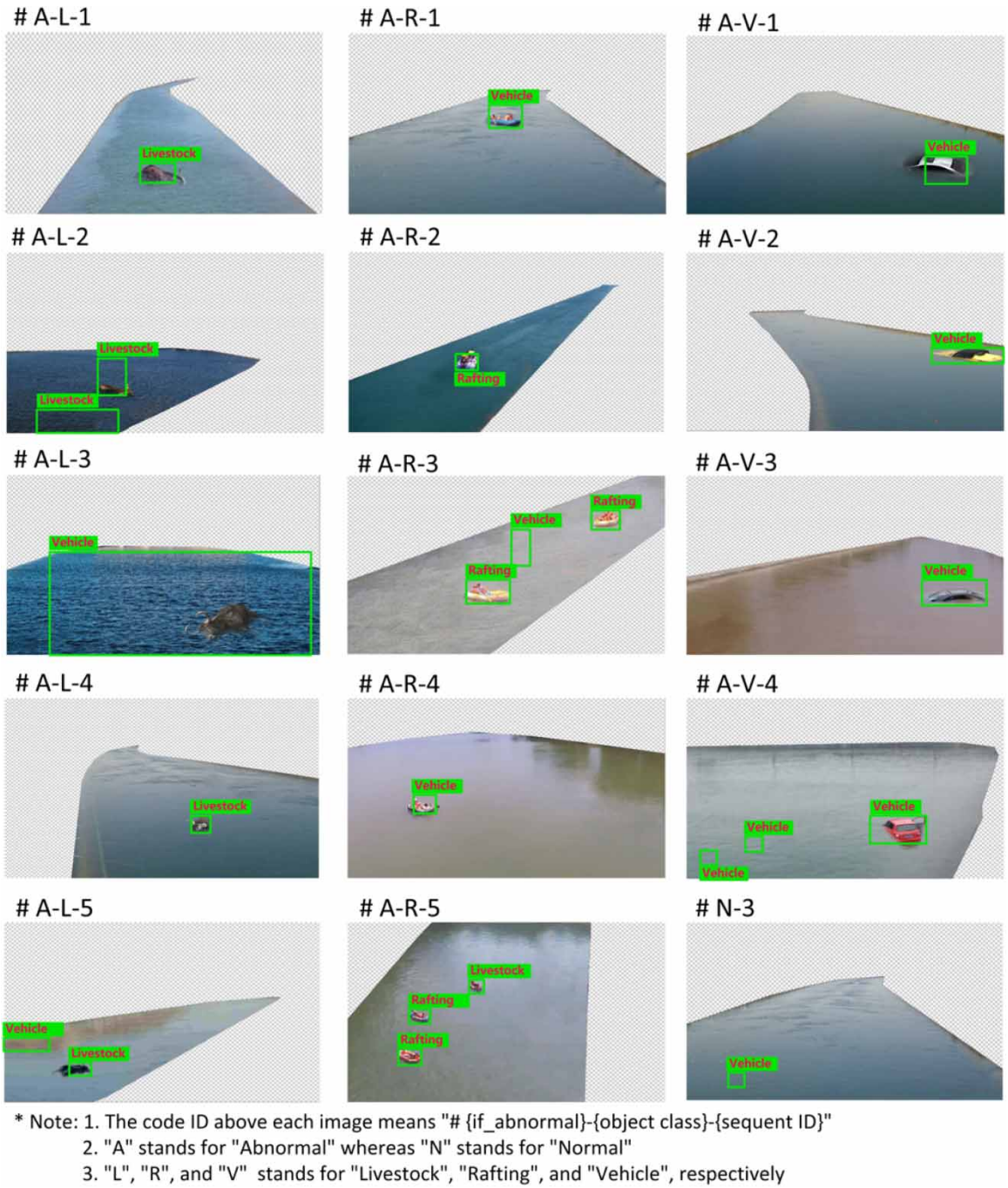


Figure 7 | Bounding boxes and classification results of detected foreign instances in the test set.

to cluster detected abnormal regions into individual instances based on the spatial distance. The proposed approach is applicable to the detection of other objects of interest in water as well. For example, our approach and findings would be of value to scenarios such as search and rescue operations at sea/desert and garbage salvage in the municipal water area. This study mainly focused on three typical foreign objects during the operation of water channels (i.e., livestock, vehicle, and rafting), which, certainly, are not exhaustive. However, with the proposed approach, the same paradigm of 'data preprocessing–anomaly detection–instance clustering–instance classification' can be extended to detect, extract, and classify other types of foreign objects once relevant training data are available.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

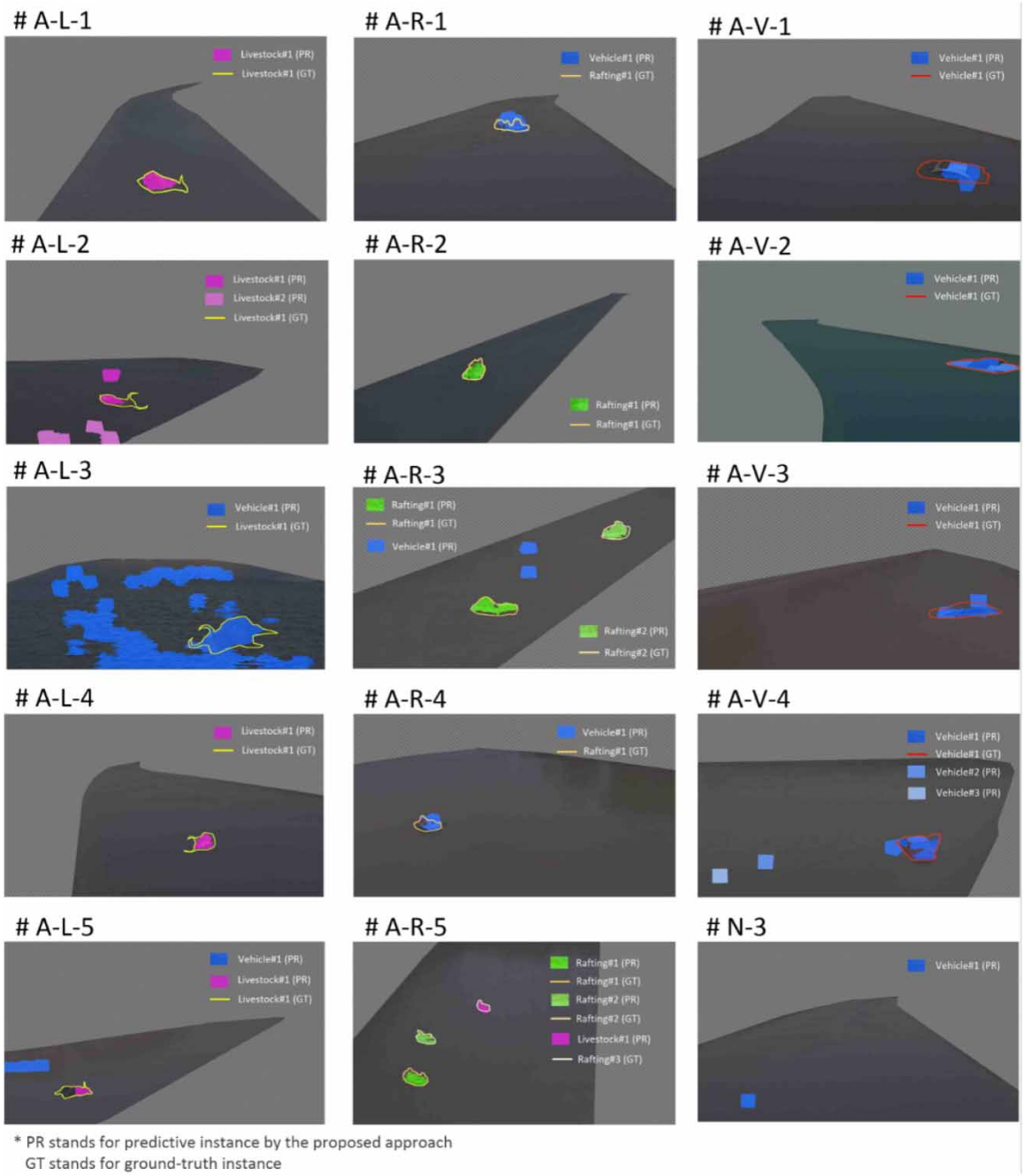


Figure 8 | Pixelwise segmentation results of the detected instances in the test set.

The proposed approach provides a promising alternative to current manual inspection for foreign object detection. With UAV applied, the condition of the entire span of a water channel can be recorded by an airborne camera in a timely manner. The recorded aerial photographs can be automatically processed by our approach to identify if an intrusion incident has occurred, and if so, the number and categories of foreign objects can also be automatically extracted for the reference of decision-makers. Another advantage of the proposed approach is that it provides geometric information of the foreign objects' boundary. Based on the geolocation of the UAV and the rationale of photogrammetry, such information can be used to estimate the actual dimensions of detected objects, which is of importance for decision-makers to evaluate the scale of intrusion. If integrated into a real-life process, the approach can be expected to streamline the current practice of water quality management in water supply channel, and thus dramatically improve the efficiency and timeliness.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Limitations and future work

Despite the efficacy demonstrated by the experimental studies, future research is recommended to address the following limitations. First, to avoid the adverse influence of irrelevant objects, this study assumes the images being processed only contain the ROI, i.e., water body. Even though the process of ROI extraction can be automated by the method introduced in Chen *et al.* (2019), it is difficult to guarantee that such extraction can completely remove the irrelevant background. As a result, in future study, it is necessary to comprehensively evaluate the performance by integrating the previous ROI extraction method and the proposed foreign object detection approach in a unified case study.

Second, criticism might be raised concerning the generalizability of the model, which was trained on a relatively small dataset. Even though relevant measures have been taken to ensure the collected dataset encompasses as many variations as possible (see Section S4 of the Supplementary Material), the model's robustness still requires further evaluation and improvement. One possible way to do that is to continuously calibrate and update the model as more and more real-life data are collected in production.

Third, while the model successfully detected most of the foreign objects, it also issued several false alarms. Such false alarms can cause unnecessary waste of resources for emergency response. Measures can be taken to minimize the negative effects of false alarms from both technical and managerial perspectives. From technical perspective, more features (e.g., color histogram) can be incorporated to the SVM model to further facilitate the performance in discriminating between water body and foreign objects. From a managerial perspective, specific personnel can be assigned to double-check if the issued alarms are correct before further reactive measures are taken.

CONCLUSION

Timely detection of foreign objects in water channel and acquiring their relevant information are of great importance to ensure water supply safety. Current practice of foreign object detection relies on manual onsite patrolling, which is inefficient, laborious, and time-consuming. UAVs can serve as a promising alternative by providing time-sensitive and actional aerial images, from which information of intrusion can be automatically extracted. This paper presents a novel approach to detecting, extracting, and classifying foreign objects from aerial images for hazard management of water channels. The collected aerial images are first preprocessed by a superpixel segmentation algorithm. A SVM is trained to detect abnormal regions, which are then clustered into individual instances based on the spatial distance. Finally, a hierarchical voting mechanism is devised to classify the extracted foreign instances. Experimental studies were implemented to evaluate the proposed approach step by step. With a small collection of 121 training samples, a well-performed model was successfully constructed, which can detect, extract, and classify three types of foreign objects (i.e., livestock, rafting, and vehicle) in the test set with a robust performance. The proposed foreign object detection approach, integrated with UAVs, is expected to raise early and comprehensive awareness of the intrusion incidents in water channels, thus providing technical tools to ensure the water supply safety.

ACKNOWLEDGEMENTS

This research was supported by the National Key Research and Development Program of China (No. 2018YFC0406903) and the National Natural Science Foundation of China (No. 51979189).

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. & Süsstrunk, S. 2012 SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**, 2274–2282. https://doi.org/10.1109/TPAMI.2012.120.

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Esesn, B. C., Awwal, A. A. S. & Asari, V. K. 2018 The history began from AlexNet: a comprehensive survey on deep learning approaches. *arXiv preprint*. arXiv:1803.01164.

Azar, E. R. & McCabe, B. 2012 Automated visual recognition of dump trucks in construction videos. *Journal of Computing in Civil Engineering* **26**, 769–781. doi:10.1061/(Asce)Cp.1943-5487.0000179.

Bhola, R., Krishna, N. H., Ramesh, K. N., Senthilnath, J. & Anand, G. 2018 Detection of the power lines in UAV remote sensed images using spectral-spatial methods. *Journal of Environmental Management* **206**, 1233–1242. https://doi.org/10.1016/j.jenvman.2017.09.036.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Cai, S., Ma, Z., Skibniewski, M. J. & Bao, S. 2019 Construction automation and robotics for high-rise buildings over the past decades: a comprehensive review. *Advanced Engineering Informatics* **42**, 100989. https://doi.org/10.1016/j.aei.2019.100989.

Carlier, J., Davis, E., Ruas, S., Byrne, D., Caffrey, J. M., Coughlan, N. E., Dick, J. T. A. & Lucy, F. E. 2020 Using open-source software and digital imagery to efficiently and objectively quantify cover density of an invasive alien plant species. *Journal of Environmental Management* **266**, 110519. https://doi.org/10.1016/j.jenvman.2020.110519.

Cha, Y.-J., Choi, W. & Buyukozturk, O. 2017 Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering* **32**, 361–378. doi:10.1111/mice.12263.

Chen, J. & Liu, D. 2021 Bottom-up image detection of water channel slope damages based on superpixel segmentation and support vector machine. *Advanced Engineering Informatics* **47**, 101205. https://doi.org/10.1016/j.aei.2020.101205.

Chen, J., Liu, D., Li, S. & Hu, D. 2019 Registering georeferenced photos to a building information model to extract structures of interest. *Advanced Engineering Informatics* **42**, 100937. https://doi.org/10.1016/j.aei.2019.100937.

Chen, J., Liu, D. & Li, X. 2021 Extracting water channels from aerial videos based on image-to-BIM registration and spatio-temporal continuity. *Automation in Construction* **132**, 103970. https://doi.org/10.1016/j.autcon.2021.103970.

civilServant-666. 2021 Foreign_Obj_in_Water_Channel. Available from: https://github.com/civilServant-666/Foreign_obj_in_water_channel (accessed 12 November 2021).

Deng, J., Yin, H., Kong, F., Chen, J., Dronova, I. & Pu, Y. 2020 Determination of runoff response to variation in overland flow area by flow routes using UAV imagery. *Journal of Environmental Management* **265**, 109868. https://doi.org/10.1016/j.jenvman.2019.109868.

Desh Deep. 2019 MP: Car Sinking in River after Accident. Available from: https://timesofindia.indiatimes.com/city/bhopal/mp-car-sinking-in-river-after-accident-father-throws-6-month-old-like-a-ball/articleshow/71809332.cms (accessed 14 November 2021).

Du, Y., Song, K., Liu, G., Wen, Z., Fang, C., Shang, Y., Zhao, F., Wang, Q., Du, J. & Zhang, B. 2020 Quantifying total suspended matter (TSM) in waters using Landsat images during 1984–2018 across the Songnen Plain, Northeast China. *Journal of Environmental Management* **262**, 110334. https://doi.org/10.1016/j.jenvman.2020.110334.

Farooquee, N. A., Budal, T. K. & Maikhuri, R. 2008 Environmental and socio-cultural impacts of river rafting and camping on Ganga in Uttarakhand Himalaya. *Current Science*, 587–594.

Jia, J. 2016 A technical review of hydro-project development in China. *Engineering* **2**, 302–312. https://doi.org/10.1016/J.ENG.2016.03.008.

Jones, D. I. 2007 An Experimental Power Pick-Up Mechanism for an Electrically Driven UAV.

Kim, I.-H., Jeon, H., Baek, S.-C., Hong, W.-H. & Jung, H.-J. 2018 Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. *Sensors* **18**. doi:10.3390/s18061881.

Kim, D., Liu, M., Lee, S. & Kamat, V. R. 2019 Remote proximity monitoring between mobile construction resources using camera-mounted UAVs. *Automation in Construction* **99**, 168–182. https://doi.org/10.1016/j.autcon.2018.12.014.

Lei, L. 2019 *Object Detection for Floating Objects Based on Deep Learning*. Guangxi University.

Li, Z. & Yang, X. 2020 Fusion of high- and medium-resolution optical remote sensing imagery and GlobeLand30 products for the automated detection of intra-urban surface water. *Remote Sensing* **12**, 4037.

Lin, Y. 2020 *Research on Foreign Object Detection on Urban River Surface Based on Machine Vision*. Zhejiang University.

Liu, D., Chen, J., Hu, D. & Zhang, Z. 2019 Dynamic BIM-augmented UAV safety inspection for water diversion project. *Computers in Industry* **108**, 163–177. https://doi.org/10.1016/j.compind.2019.03.004.

Lucey, A. & Glennon, N. 2021 Dumping of Dead Lambs in Kerry River Happening on 'Consistent Basis'. Available from: https://www.irishexaminer.com/news/munster/arid-40241426.html (accessed 14 November 2021).

Melo, R. R. S. d., Costa, D. B., Álvares, J. S. & Irizarry, J. 2017 Applicability of unmanned aerial system (UAS) for safety inspection on construction sites. *Safety Science* **98**, 174–185. https://doi.org/10.1016/j.ssci.2017.06.008.

Norvanchig, J. & Randhir, T. O. 2021 Simulation of ecohydrological processes influencing water supplies in the Tuul River watershed of Mongolia. *Journal of Hydroinformatics* **23**, 1130–1145. doi:10.2166/hydro.2021.056.

Ojala, T., Pietikainen, M. & Maenpaa, T. 2002 Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**, 971–987. doi:10.1109/TPAMI.2002.1017623.

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

Pietikäinen, M., Ojala, T. & Xu, Z. 2000 Rotation-invariant texture classification using feature distributions. *Pattern Recognition* **33**, 43–52. https://doi.org/10.1016/S0031-3203(99)00032-1.

Ren, X. & Malik, J. 2003 Learning a classification model for segmentation. In: *Proceedings of the Ninth IEEE International Conference on Computer Vision*, vol. 11, pp. 10–17.

Rengaraju, P., Pandian, S. R. & Lung, C. 2014 Communication networks and non-technical energy loss control system for smart grid networks. In: *2014 IEEE Innovative Smart Grid Technologies – Asia (ISGT ASIA)*, pp. 418–423.

Rezazadeh Azar, E. & McCabe, B. 2012 Part based model and spatial–temporal reasoning to recognize hydraulic excavators in construction images and videos. *Automation in Construction* **24**, 194–202. https://doi.org/10.1016/j.autcon.2012.03.003.

Schumacher, J. 2002 Surface Water Pollution from Livestock Production. ExEx.

Wang, H., Zhang, W., Hong, S., Zhuang, Y., Lin, H. & Wang, Z. 2013 Spatial evaluation of complex non-point source pollution in urban–rural watershed using fuzzy system. *Journal of Hydroinformatics* **16**, 114–129. doi:10.2166/hydro.2013.266.

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest

Uncorrected Proof

Wang, J., Sun, W., Shou, W., Wang, X., Wu, C., Chong, H.-Y., Liu, Y. & Sun, C. 2015 Integrating BIM and LiDAR for real-time construction quality control. *Journal of Intelligent & Robotic Systems* **79**, 417–432. doi:10.1007/s10846-014-0116-8.

Yan, J., Jin, J., Chen, F., Yu, G., Yin, H. & Wang, W. 2017 Urban flash flood forecast using support vector machine and numerical simulation. *Journal of Hydroinformatics* **20**, 221–231. https://doi.org/10.2166/hydro.2017.175.

Yu, Y., Guan, H., Zai, D. & Ji, Z. 2016 Rotation-and-scale-invariant airplane detection in high-resolution satellite images based on deep-Hough-forests. *ISPRS Journal of Photogrammetry and Remote Sensing* **112**, 50–64. https://doi.org/10.1016/j.isprsjprs.2015.04.014.

Zhang, L., Li, S., Loáiciga, H. A., Zhuang, Y. & Du, Y. 2015 Opportunities and challenges of interbasin water transfers: a literature review with bibliometric analysis. *Scientometrics* **105**, 279–294. doi:10.1007/s11192-015-1656-9.

Zhang, S., Liu, T. & Wang, C. 2021 Multi-source data fusion method for structural safety assessment of water diversion structures. *Journal of Hydroinformatics* **23**, 249–266. doi:10.2166/hydro.2021.154.

Zhu, P., Sun, Y., Wen, L., Feng, Y. & Hu, Q. 2020 Drone Based RGBT Vehicle Detection and Counting: A Challenge.

First received 25 August 2021; accepted in revised form 15 November 2021. Available online 10 December 2021

Downloaded from http://iwaponline.com/jh/article-pdf/doi/10.2166/hydro.2021.118/975962/jh2021118.pdf
by guest