STUDY AND APPLICATION OF IMAGE WATER LEVEL RECOGNITION CALCULATION METHOD BASED ON MASK R-CNN AND FASTER R-CNN

LIANG, L.[1] – HUANG, W.[1] – AWAN, M. I.[2] – PARVEEN, A.[2] – LI, R. P.[1] – BI, F. G.[1] – SHAO, J. H.[1*] – LIANG, X. Y.[1] – WU, C. H.[1] – LIU, Z. Q.[1*]

[1]*Guangxi Hydraulic Research Institute, Guangxi Key Laboratory of Water Engineering Materials and Structures, Nanning 530023, China*

[2]*Department of Agronomy, Sub-Campus Depalpur, University of Agriculture Faisalabad, Okara, Pakistan*

[*]*Corresponding authors*
e-mail: jinhua20211103@outlook.com, m15078812237@163.com

(Received 19th May 2023; accepted 11th Aug 2023)

Abstract. To measure the water level with a water level gauge, an automatic water level calculation method based on Region-based Convolutional Neural Networks (R-CNN) is proposed in this study. First, the water level gauge is located and the contour of water level gauge is obtained using the enhanced Mask R-CNN model approach. Secondly, using the contour of the water gauge, image processing technology is utilized to recognize and intercept the image that only contains the water gauge. Next, the scale on the water level gauge is detected and identified using the Faster R-CNN (Region-based Convolutional Neural Networks) model approach. The projection transformation method then calculates the value of the current water level based on the scale of the identified water level. The proposed approach was demonstrated by experimental findings to be relatively straightforward, to have high recognition accuracy and ability in complex interference environments in the field, and to have low cost and great application value. The findings presented in this paper can be used to help with automatic water level gauge measuring both theoretically and practically.
Keywords: *CNN, Mask R-CNN, Faster R-CNN, image recognition, water level calculation, water resources*

Introduction

Artificial intelligence technology has advanced quickly in recent years and has demonstrated a wide range of potential applications in numerous industries (Ali et al., 2023). Image-based automated measuring technology has received a lot of attention and use as big data, computer vision, and deep learning. One of the most demanding evaluations in water resource management is the effective utilization coefficient of irrigation water, and the accurate regulation and measurement of irrigation water that depends greatly on the effectiveness of automated water level metering technology (Guo et al., 2022; Fang et al., 2023). It is challenging to promote and operate automated metering systems in less developed areas due to issues with their high cost, weak anti-interference capabilities, and high maintenance costs (Yan et al., 2021). The best option in this situation is to measure and recognize the water level automatically utilizing image-based automatic measuring technology. This paper suggests a straightforward and effective technique for automatic water level measurement and recognition, which can recognize the water level in images when computational resources are limited and sample sizes are small (Zhang et al., 2019). Additionally, compared to monitoring devices like radar and ultrasound, this method's hardware equipment installation and maintenance

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

costs are lower. It only requires the use of a surveillance camera, which is very useful, particularly in the management of water resources and irrigation water in underdeveloped areas (Fang et al., 2023).

There are numerous studies currently focused on enhancing the precision and effectiveness of hydrological monitoring through the use of computer vision technology and machine learning techniques. Peng et al. (2012) devised a method for identifying waterlines that combines morphology and the “Canny” algorithm (computational theory of edge detection). However, this method necessitates good image quality, and the threshold used during edge extraction has a considerable impact on recognition accuracy. The projection approach to segment water gauge symbols recognize water gauge characters using a BP neural network, but this method is heavily influenced by environmental factors such as lighting conditions (Lin et al., 2022). This strategy, however, requires discriminating hues before recognition, which is easily influenced by the surroundings. Dou et al. (2022) proposed a water gauge reading recognition approach based on an upgraded UNet network that can recognize water surface regions and potential water gauge marker locations. The images are then classified and recognized using convolutional neural networks, and the real water gauge reading is determined at the end (Faraj et al., 2022). However, this method's implementation is relatively difficult, needing a big amount of data to train the network and a considerable recognition time (Zhang et al., 2023). Furthermore, disturbances such as illumination and water dips have a strong influence on the recognition rate (Han et al., 2021)

This paper employs an inventive target identification algorithm based on the combination of Mask R-CNN (He et al., 2017) and Faster R-CNN (Ren et al., 2017) models to find water gauges and identify water gauge scales, drawing inspiration from the research accomplishments of the aforementioned scholars. Although "waterlines" have been discussed by earlier researchers, they are not necessary for the method in this article. Instead, the final calculation of the actual water level uses the projection transformation method. It can increase the success rate and stability of recognition, even though the accuracy of identifying and calculating water levels may not necessarily improve significantly. This makes it particularly suitable for promotion and use in the fields of extensive water level monitoring and water quantity statistics.

Materials and Methods

Study Area

This study was carried out in Guangxi governorate (province), which is one of the South Region China governorates with a GPS reading of latitude = 22.815478; longitude = 108.327546 (*Figure 1*). The current study used the resource of 15 locations which applied of image recognition water level technology and the study’s images were taken in those locations.

Methods Process

The purpose of this paper is to address the issue of big disparities in target sizes and easy target loss in water gauge photos. The detection of a single target in a water gauge and more accurate water gauge edge segmentation is obtained by applying pixel-level segmentation technology based on the Mask R-CNN algorithm. Mask R-CNN is a descendant of Faster R-CNN. Thus, Faster R-CNN was improved, and mask branches

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

were added to provide pixel-level classification and instance segmentation. In this study, ResNet101 (He et al., 2016) is used in conjunction with FPN (Lin et al., 2016) for multi-scale feature extraction, RPN (Region Proposal Network) network for proposed region generation, Head branch for suggested region classification, bounding box regression, and pixel-level classification. Additionally, target scales of the same size are swiftly identified and segmented using the Faster R-CNN algorithm. Although the algorithm in this article includes significant redundancy, it can recognize water level gauges and calculate water levels with greater accuracy with less computational resources and sample data.

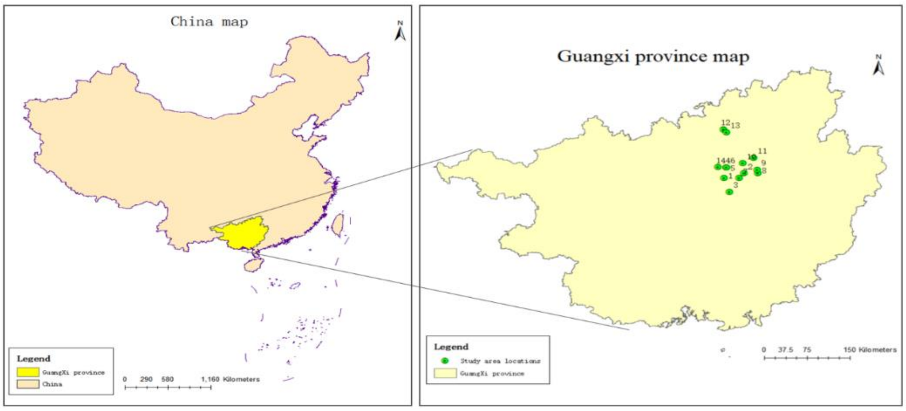


Figure 1. *Study area locations map*

The recognition and positioning of water level gauge targets is the cornerstone of water level computation that plays a critical role in recognizing water levels and is also an important stage in water level image recognition study. The issue lies in precisely finding the position of the water level gauge on the water surface in the image, which directly influences recognition accuracy. This method is vulnerable to factors such as the reflection of the water level gauge and the occlusion of the cover. The essential steps of the regional convolutional neural network-based water level recognition system investigated in this paper are as follows: ① image preprocessing, ② model training, ③ water gauge target prediction, ④ water gauge positioning (Mask R-CNN model), ⑤ water gauge clipping, ⑥ scale positioning (Faster R-CNN model), ⑦ scale character recognition, ⑧ water level calculation, as shown in the following figure (*Figure 2.*).

The Mask R-CNN Model

The Principle of Recognition

The primary portion of the water gauge is recognized using the Mask R-CNN technique, which can segment the image down to the pixel level but operates at a somewhat slow speed. The employment of Mask R-CNN can more correctly segment the gauge's edge for the recognition of a single target in a water gauge and can also handle targets of various sizes. Fundamentally, Faster R-CNN and FCN are combined in the Mask R-CNN algorithm. ResNet and the Functional Pyramid Network (FPN) are combined in this algorithm's backbone feature extraction network. ResNet (Residual Network) is a residual network that uses convolution to extract image features. However, utilizing the ResNet network to extract features alone will not be sufficient to meet the

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

needs of instance segmentation if the size of the target is randomly changing, such as the size of the targets contained in the image being varied. In order to optimize the feature extraction network and take multi-scale difficulties into account while extracting image features, FPN (Feature Pyramid Networks) was developed. To achieve feature fusion and enhance feature extraction capabilities, FPN employs two pyramid networks, top-down and bottom-up. Advanced features from the first pyramid can be extracted by the second pyramid and sent down into lower layers, allowing for full integration of features from various levels. As a result, the ResNet50+FPN backbone was utilized in this experiment. Below *Figure 3.* depicts the construction of the Mask R-CNN algorithm model.

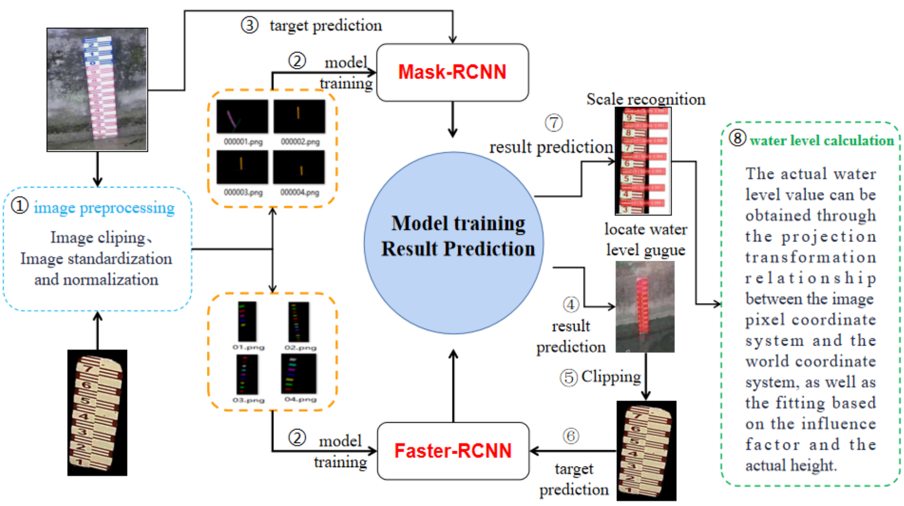


***Figure 2.** Process involved in water level calculation method based on Mask-RCNN*

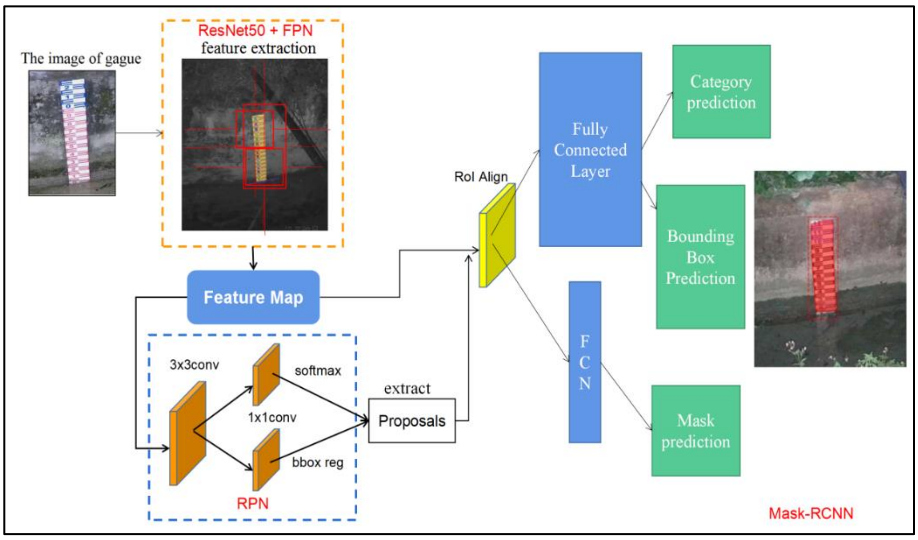


***Figure 3.** Construction of the Mask-RCNN algorithm model*

Model Training

We must first realize that Mask R-CNN is built on Faster R-CNN in order to apply it. In order to accomplish the function of instance segmentation (target detection in the

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

image), a mask layer was added to the initial infrastructure. For each image sample's RoI (Region of Interest), Mask R-CNN will produce a binary mask. For each RoI, a multitask Loss function is defined during training. As we are aware, the loss function typically consists of three parts. The following equation (*Eq.1*) illustrates this:

$$\mathrm{Loss} = \mathrm{L}_{\mathrm{cls}} + \mathrm{L}_{\mathrm{reg}} + \mathrm{L}_{\mathrm{mask}} \tag{Eq.1}$$

wherein,

$$\mathrm{L}_{\mathrm{cls}} = \frac{1}{\mathrm{N}_{\mathrm{cls}}} \sum_{i} \mathrm{l}_{\mathrm{cls}}(p_i, p_i^{*}) \tag{Eq.2}$$

$$\mathrm{L}_{\mathrm{reg}} = \lambda \frac{1}{\mathrm{N}_{\mathrm{reg}}} \sum_{i} \mathrm{p}_{\mathrm{i}}^{*} \mathrm{l}_{reg}(t_i, t_i^{*}) \tag{Eq.3}$$

$$\mathrm{L}_{\mathrm{mask}} = \frac{1}{m^2} \sum_{i}^{k} (l^k) \sum_{l}^{m^2} [-y * \log(\sigma(x)) - (l-y) * \log(l-\sigma(x))] \tag{Eq.4}$$

The classification loss (*Eq.1*), regression box loss (*Eq.2*), and mask loss (*Eq.3*) calculations are shown above. The Loss function's primary job is to determine how much the network output result differs from the actual value. The pertinent information contains the particular meaning of each formula; therefore it is not necessary to go into detail here. We employ the following formula (*Eq.5*) for straightforward expression to aid comprehension:

$$\mathrm{Loss} = \frac{1}{2} \sum_{i}^{\mathrm{n}} (\mathrm{x}_i - y_i)^2 \tag{Eq.5}$$

The loss relationship curve can be created using equation (*Eq.5*), where yi is the expected value and xi is the network output. The loss function is minimized using the Gradient descent approach during the training of the Mask R-CNN model. There are 450 samples in this training set, and the following key model training parameters are set:

IMAGE_MIN_DIM = 480,
IMAGE_MAX_DIM = 1024,
RPN_ANCHOR_SCALES = (32 * 6, 64 * 12, 128 * 24, 256 * 36, 512 * 48) ,
TRAIN_ROIS_PER_IMAGE = 10,
STEPS_PER_EPOCH = 10,
VALIDATION_STEPS = 10,
EPOCH=30.

We can tell whether the model was correctly trained, whether it was adequately trained, whether it was overfitted, and whether the initialization weight was set appropriately using the Loss function. In order to gauge how well the model presented itself during

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

training, we kept track of changes in the loss function and the accuracy of the training set and verification set. last but not least, we plotted the trend of the loss function value loss with the number of iterations in accordance with the model training outcomes, as shown in the accompanying image (*Figure 4.*).

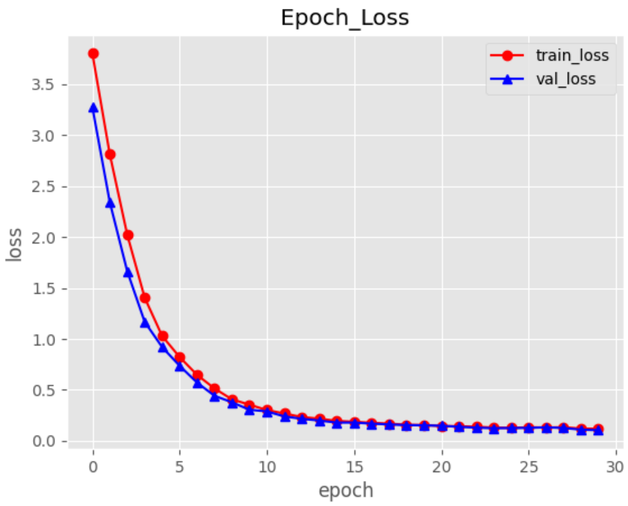


Figure 4. *Mask R-CNN Model Training (38s/step, 380s/epoch)*

Model Verification

ROI Align replaces ROI Pooling in Mask R-CNN. By using pooling, ROI Pooling adjusts the feature map's size so that it is the same size as the original ROI. In order to effectively prevent error dislocation and maintain consistency between the size of the original image and the feature map, ROI Align uses bilinear interpolation. This allows for the acquisition of more precise regional location data. The image samples are real-time collected photos of the 15 stations with security cameras stated in the earlier "study area" section. This model uses 450 image samples as the training set and 750 image samples as the validation set. It can be shown that this model still has decent recognition accuracy even with a minimal number of samples and training times. The validation outcomes for this model are shown in *Table 1* below.

Table 1. *The result of Mask R-CNN model verification*

Number of Samples Trained	Number of Samples Tested	Recognition Accuracy
450	750	92.19%

The Faster R-CNN Model

The Principle of Recognition

We used the Faster R-CNN algorithm to identify the water gauge scale because it is a little target and the Mask R-CNN method does not perform well in terms of detection performance or speed when dealing with a large number of small targets. Four components make up the Faster R-CNN's core structure: feature extraction, RPN, proposal layer, and ROI pooling. The RPN part is a brand-new structure of Faster R-CNN, which determines the estimated position of the target from the Feature Map during

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

network training; The feature extraction part uses convolution and pooling methods to extract Feature Map from the original picture; The Proposal Layer portion makes use of the roughly determined positions by RPN for additional training to determine more exact positions; the ROI Pooling section makes use of the precisely determined positions from earlier to extract targets for classification from the Feature Map. The Faster R-CNN method outperforms and detects small targets (pictures smaller than 32x32 pixels) more quickly than the Mask R-CNN algorithm (*Figures 5, 6*).

Figure 5. *The effect of water gauge target recognition and edge extraction*

Model Training

As well known, generally, the Loss function Loss (*Eq.6*) of the Faster R-CNN model consists of two parts: classified loss (*Eq.7*) and regression box loss (*Eq.8*), as shown below:

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

$$\mathrm{Loss} = \mathrm{L}_{\mathrm{cls}} + \mathrm{L}_{\mathrm{reg}} \tag{Eq.6}$$

wherein,

$$\mathrm{L}_{\mathrm{cls}} = \frac{1}{\mathrm{N}_{\mathrm{cls}}} \sum_{i} \mathrm{l}_{\mathrm{cls}}(p_i, p_i^*) \tag{Eq.7}$$

$$\mathrm{L}_{\mathrm{reg}} = \lambda \frac{1}{\mathrm{N}_{\mathrm{reg}}} \sum_{i} \mathrm{p}_{\mathrm{i}}^* \mathrm{l}_{reg}(t_i, t_i^*) \tag{Eq.8}$$

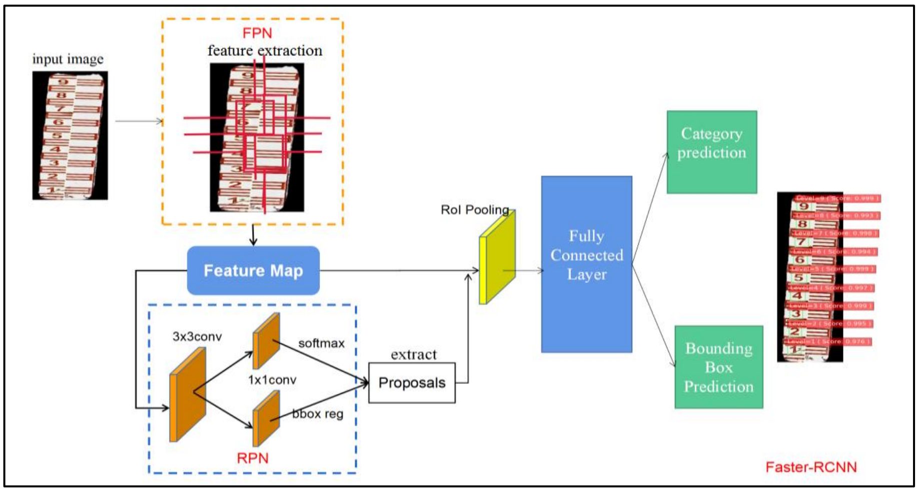


Figure 6. *Faster R-CNN model structure*

The Loss function is mainly used to evaluate the difference between the network output result and the actual value. Similarly, we can use the following formula (*Eq.8*) to simply express it:

$$\mathrm{Loss} = \frac{1}{2} \sum_{i}^{\mathrm{n}} (\mathrm{x}_i - y_i)^2 \tag{Eq.9}$$

Network output (x_i) and expected value (y_i) are both included in equation (*Eq.8*). It is possible to draw a Loss relationship curve as a result. 350 samples make up the Faster R-CNN model training set, and these key parameters are used to train the model:

weight_decay＝0.0005,

learning_rate＝0.001,

momentum＝0.9,

gamma＝0.1,

batch_size＝256,

max_iters＝40000,

step_size＝30000.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

Finally, according to the model training results, the change trend of the Loss function value Loss with the number of iterations is shown in the figure (*Figure 7.*) below:

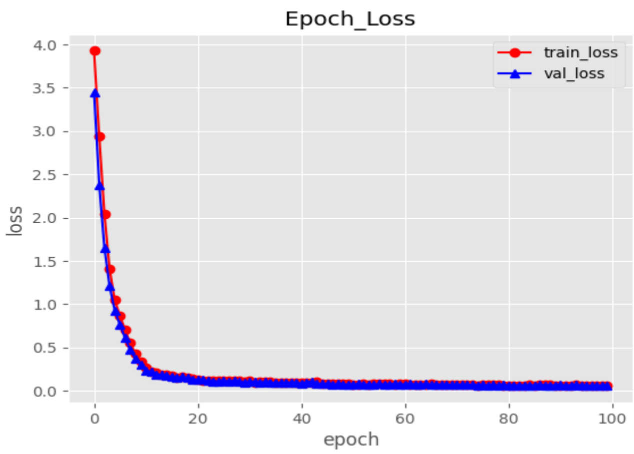


Figure 7. *Faster R-CNN model training (38s/step, 380s/epoch)*

Model Verification

Similarly, this model employs 450 picture samples for the training set and 750 for the validation set (the image samples are taken from the real-time collected photos of the 15 stations outfitted with security cameras stated in the previous "study area" section). As can be observed, the Faster R-CNN algorithm allows for quick segmentation and recognition for targets of the same size. This model also has good recognition accuracy even with a limited amount of samples and training sessions. Below, in *Table 2*, are the validation outcomes for this model.

Water Level Calculation

The Calculation Principle

We can pinpoint the water gauge's location and recognize its scale using the first two models. The link between the water gauge and scale in the pixel coordinate system of the image and the world coordinate system is projected and transformed to determine the real water level value. The accompanying figure below illustrates the calculation's basic principle (*Figure 8*).

According to the *Figure 8*, the relationship equation between the water gauge scale and the actual water level can be obtained:

$$\frac{\Delta H_0^{'}}{\Delta H^{'}} = \frac{\Delta H_0}{\Delta H}\kappa \tag{Eq.10}$$

wherein,

$\Delta H_0^{'}$ is the height of the character scale in the image; $\Delta H^{'}$ is the height of the gauge scale reading to the water surface line in the image; ΔH_0 is the actual height of the gauge scale reading; ΔH is the height of the actual gauge scale reading to the water

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

surface line; H is the actual water level; and κ is the projection transformation coefficient between the pixel coordinate system and the world coordinate system. (related to camera internal and external parameters, lens distortion parameters, and other influencing factors).

$$H = N - \Delta H = N - \frac{\Delta H_0 * \Delta H'}{\Delta H_0'}\kappa \qquad \text{(Eq.11)}$$

Table 2. *The result of Mask R-CNN model verification*

Number of Samples Trained	Number of Samples Tested	Recognition Accuracy
450	750	92.6%
	Recognition Effect of Water Gauge Scale	

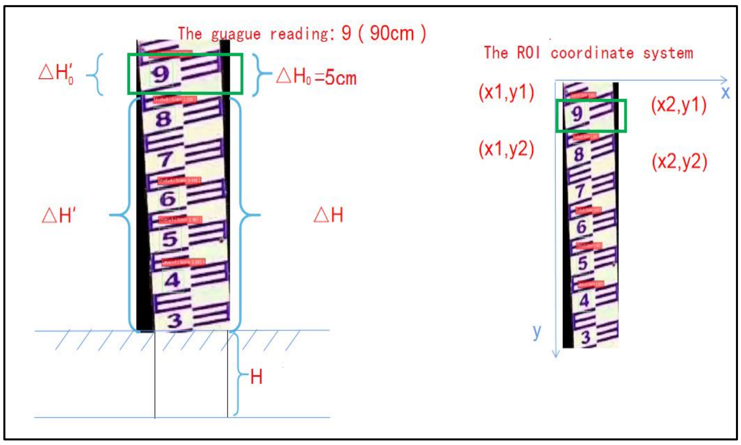


Figure 8. *The Calculation principle*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

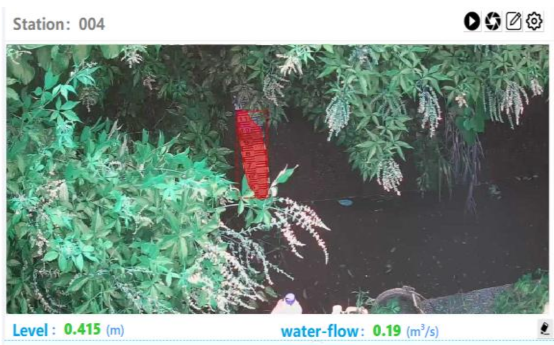
Result Analysis

100 image samples were used as the verification set for the verification of the calculation of water level recognition based on the scale of the water level gauge. The image samples came from the real-time captured images of the 15 stations equipped with monitoring cameras mentioned in the earlier "study area" section. *Table 3* below displays the actual findings of the water level verification. Even with fewer samples and shorter training periods, this model has a high computation success rate of 98%. However, the calculation error range of the final water level value is quite significant (±2.5 cm) due to the short number of samples and model training time-frames, as well as the uncertainty of the " κ " factor raised in equation "*Eq.11*".

Table 3. *The result of Mask R-CNN model verification*

Number of Samples Tested	Calculated Success Rate	Calculated Error Range
100	98%	±2.5cm

Actual water level calculation results



Discussion

Comparison of Cost Performance

Traditional methods of measuring water levels rely on visual readings, which are labor-intensive and not real-time, making it impossible to capture data at any moment and guaranteeing the continuation of water conservancy measurements (Fang et al., 2023). However, the majority of sensors used for automatic collection use pressure, float mechanical coding, radar, ultrasonic, and other measuring techniques (Janai et al., 2020). These measurement techniques are expensive to implement, difficult to maintain, have a

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

substantial environmental impact, are exposed to wind and rain for extended periods of time outdoors, are easily damaged, and have other drawbacks (Peng et al., 2019).

A straightforward picture capturing device is all that is needed in terms of hardware for the water level gauge in this approach. It has the ability to automatically measure water level using a standard water gauge and to keep an eye on localized video (Shorten and Khoshgoftaar, 2019; Liang et al., 2022). This method's hardware is simpler and less expensive to install and maintain than monitoring devices like radar and ultrasound because it simply requires monitoring cameras. It may also be used for comparing and calibrating water level sensors. It can be widely employed in a variety of water level measurement settings. Additionally, it can be used to calibrate and compare water level sensors (Xu et al., 2020; Lin et al., 2022). This technology is particularly well suited for promotion and usage in the field of extensive water level monitoring and water volume statistics (without the need for too exact computations), despite having low measurement accuracy but high cost-effectiveness (*Table 4*).

Table 4. *The water level gauge products*

Water Level Gauge Products	Main Hardware	Measurement Error	Price (USD)
Derivatives Of This Method	5 million pixels Camera	±2cm	83.00
Radar Water Level Gauge	Radar water level gauge, RTU	±0.05cm	1500.00
Ultrasonic Water Level Gauge	Ultrasonic water level gauge, RTU	±0.02cm	1800.00
Float Type Water Level Gauge	Float type water level gauge, RTU	±0.1cm	1100.00
Other Water Level Gauges	-	-	>200.00

It should be noted that this method's cost for derivative water level gauges is significantly cheaper than its cost for other kinds of water level gauges. Other water level gauges also require computer support to be applied to pertinent business systems, even though this method uses computers to execute complex models. Therefore, there is no need to talk about how much computers cost in terms of hardware here

Comparison of Effectiveness

The automatic water level measuring algorithm shown in this study is based on Mask R-CNN and Faster R-CNN and has three components: locating the water gauge, reading the scale characters, and calculating the water level (Yuan et al., 2019; Pobar and Ivašić, 2019). First, using the water gauge's feature information, the Mask R-CNN model is utilized to precisely find the edge of the water gauge. Second, character recognition on the water gauge was performed using Faster R-CNN, and segmented character recognition was accomplished using a CNN. The measuring results that followed were based on the recognition results. A quadratic equation fitting based on the influence factor and real height is utilized to establish measuring standards based on the pixel height of characters.

During the training process of multi-target images with different sizes using the Mask R-CNN model algorithm, the loss function convergence is not optimal or does not converge (see *Figure 9. and Figure 10.*). This also confirms that the previously mentioned Mask R-CNN model algorithm is susceptible to the problem of target loss when dealing with the size difference between different objects in the image (Kang et al., 2022; Rocha et al., 2023). However, the faster R-CNN model by itself is unable to precisely segment the water gauge's boundary. It should be emphasized that the fundamental benefit of this approach is that, even with limited computer resources, fewer image samples, and short model training timeframes, it still achieves high recognition success rates.

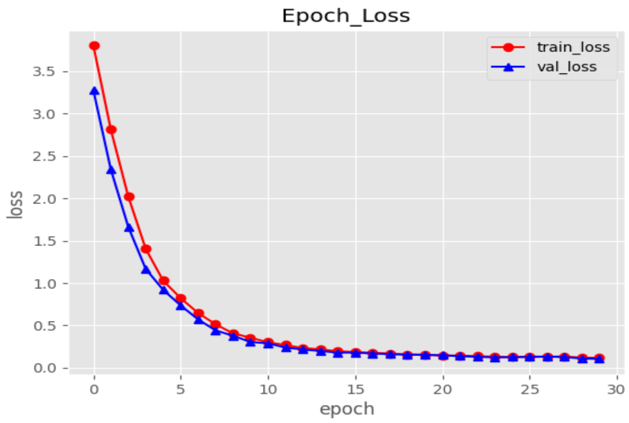


***Figure 9.** Single objective training loss (38s/step,380s/epoch)*

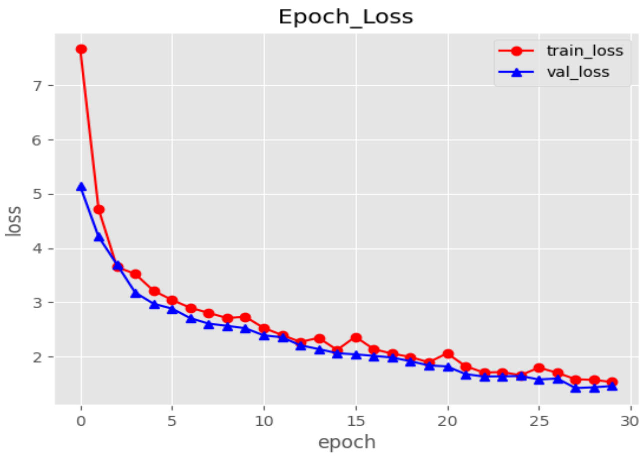


***Figure 10.** Multi-objective training loss (348s/step, 3480s/epoch)*

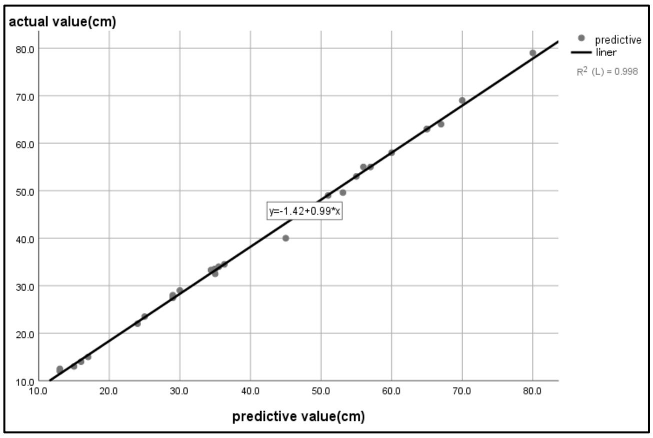


***Figure 11.** Regression analysis of actual value and predictive value*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

Conclusion

The algorithm for measuring water gauges that is suggested in this paper is based on image recognition. An intelligent measuring facility based on picture recognition is created by combining cutting-edge image processing technology with conventional water gauge measurement techniques. The method for recognizing the water level determines the water level after using image recognition and analysis algorithms. The ability to remotely monitor the water level online thanks to the real-time video camera is quite practical. With its low cost and straightforward facilities, it may be widely employed in a variety of water level measurement situations where water gauges can be used. Water resource monitoring and measuring can be "visible, accurate, and affordable" for systems that already have cameras installed by completing the restoration without adding too many facilities. Of course, the current version of this system still has a lot of flaws (in terms of accuracy, effectiveness, etc.). The key issue is that the irrigation area administrator needs to routinely clean and maintain the location because the system won't be able to accurately detect the water level if the camera or water gauge are blocked by other objects. Future work will include more analysis and improvement, as well as extensive interaction with companies that deal with water conservation.

Acknowledgements. This research was funded by Guangxi Key R&D program (Guike AB22035057) and Guangxi Key R&D program (Guike AB23026021); Beijing River Water Conservancy Development Foundation (Water conservancy youth talent development support project); Guangxi Water Conservancy Science and Technology Promotion Project (SK-2022-008 and SK2021-3- 12); Guangxi Key Laboratory of Water Engineering Materials and Structures fund program.

REFERENCES

[1] Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M. A. A., Dwivedi, Y. K. (2023): A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. – Journal of Innovative and Knowledge 8(1): 100333.

[2] Dou, G., Chen, R., Han, C., Liu, Z., Liu, J. (2022); Research on water-level recognition method based on image processing and convolutional neural networks. – Water 14(12): 1890.

[3] Fang, S., Zhang, B., Hu, J. (2023): Improved mask R-CNN multi-target detection and segmentation for autonomous driving in complex scenes. – Sensors 23(8): 3853.

[4] Faraj, R. H., Azad, A. M., Khalid, M. O., Hemn, U. A. (2022): Soft computing techniques to predict the compressive strength of green self-compacting concrete incorporating recycled plastic aggregates and industrial waste ashes. – Clean Technologies and Environment Policy 24(7): 2253-2281.

[5] Guo, M. H., Xu, T. X., Liu, J. J., Liu, Z. N., Jiang, P. T., Mu, T. J., Zhang, S. H., Martin, R. R., Cheng, M. M., Hu, S. M. (2022): Attention mechanisms in computer vision: A survey. – Computational Visual Media 8: 331-368.

[6] Han, Q., Yin, Q., Zheng, X., Chen, Z. (2021): Remote sensing image building detection method based on Mask R-CNN. – Complex and Intelligent Systems 4: 1-9.

[7] He, K., Zhang, X., Ren, S., Sun, J. (2016): Deep residual learning for image recognition. – In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778.

[8] He, K., Gkioxari, G., Piotr, D. (2017): Mask R-CNN. – IEEE Transactions on Pattern Analysis and Machine Intelligence 20(3): 281-293.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053

[9] Janai, J., Güney, F., Behl, A., Geiger, A. (2020): Computer vision for autonomous vehicles: Problems, datasets and state of the art. – Foundations and Trends in Computer Graphics and Vision 12: 1-308.
[10] Kang, K. S., Cho, Y. W., Jin, K. H., Kim, Y. B., Ryu, H. G. (2022): Application of one-stage instance segmentation with weather conditions in surveillance cameras at construction sites. – Automation in Construction 133: 104034.
[11] Liang, Y., Liao, Y., Li, S., Wu, W., Qiu, T., Zhang, W. (2022): Research on water meter reading recognition based on deep learning. – Scientific Reports 12(1): 12861.
[12] Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. (2017): Feature pyramid networks for object detection. – In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2117-2125.
[13] Lin, P., Zhang, H., Zhao, F., Wang, X., Liu, H., Chen, Y. (2022): Boosted Mask R-CNN algorithm for accurately detecting strawberry plant canopies in the fields from low-altitude drone images. – Food Science and Technology 42: e95922.
[14] Lin, F., Gan, L., Jin, Q., You, A., Hua, L. (2022): Water quality measurement and modelling based on deep learning techniques: case study for the parameter of Secchi disk. – Sensors 22(14): 5399.
[15] Peng, J., Ran, X. (2012): Method for detecting ship waterlines using image processing. – China Water Transport 6: 2.
[16] Peng, Y., Liu, X., Shen, C., Huang, H., Zhao, D., Cao, H., Guo, X. (2019): An improved optical flow algorithm based on mask-R-CNN and K-means for velocity calculation. – Applied Science 9: 2808.
[17] Pobar, M., Ivašić, K. M. (2019): Detection of the leading player in handball scenes using Mask R-CNN and STIPS. – In Eleventh International Conference on Machine Vision 11041: 501-508.
[18] Ren, S., He, K., Girshick, R., Sun, J. (2015): Faster R-CNN: Towards real-time object detection with region proposal networks. – Advances in Neural Information Processing Systems 28: 12-21.
[19] Rocha, D., Alves, J., Lopes, V., Teixeira, J. P., Fernandes, P. A., Costa, M., Morais, M., Salomé, P. M. (2023): Multidefect detection tool for large-scale pv plants: segmentation and classification. – IEEE Journal of Photovoltaics 13(2): 291-295.
[20] Shorten, C., Khoshgoftaar, T. M. (2019): A survey on image data augmentation for deep learning. – Journal of Big Data 6(1): 1-48.
[21] Xu, B., Wang, W., Falzon, G., Kwan, P., Guo, L., Chen, G., Tait, A., Schneider, D. (2020): Automated cattle counting using Mask R-CNN in quadcopter vision system. – Computers and Electronics in Agriculture 171: 105300.
[22] Yan, D., Li, G., Li, X., Zhang, H., Lei, H., Lu, K., Cheng, M., Zhu, F. (2021): An improved faster R-CNN method to detect tailings ponds from high-resolution remote sensing images. – Remote Sensing 13(11): 2052.
[23] Yuan, B., Li, Y., Jiang, F., Xu, X., Guo, Y., Zhao, J., Zhang, D., Guo, J., Shen, X. (2019): MU R-CNN: A two-dimensional code instance segmentation network based on deep learning. – Future Internet 11(9): 197.
[24] Zhang, L., Zhang, Y., Zhang, Z., Shen, J., Wang, H. (2019): Real-time water surface object detection based on improved faster R-CNN. – Sensors 19(16): 3523.
[25] Zhang, H., Li, F., Xu, H., Huang, S., Liu, S., Ni, L. M., Zhang, L., Former, M. P. (2023): Mask-piloted transformer for image segmentation. – arXiv 2023: 0733.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2106_50395053