

Article

Dynamic Multi-Parameter Sensing Technology for Ecological Flows Based on the Improved DSC-YOLOv8n Model

Jun Yu ¹, Yongsheng Li ², Ting Wang ², Peipei Zhang ^{3,*}, Wenlong Jiang ¹ and Lei Xing ¹

¹ Yellow River Engineering Consulting Co., Ltd., Zhengzhou 450003, China; yujun@yrec.cn (J.Y.); jiangwl@yrec.cn (W.J.); xinglei@yrec.cn (L.X.)

² Yunhe (Henan) Information Technology Co., Ltd., Zhengzhou 450003, China; liyongsheng411503@163.com (Y.L.); 18037777565@163.com (T.W.)

³ College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, China

* Correspondence: zhangpeipei@hhu.edu.cn

Abstract

Ecological flow management is important for maintaining ecosystem stability and promoting sustainable development. Dynamic ecological flow regulation depends on precise real-time monitoring of water levels and flow velocities. To address challenges in ecological flow monitoring, including maintenance difficulties and insufficient accuracy, an improved DSC-YOLOv8n-seg model is proposed for dynamic multi-parameter sensing, achieving more efficient object detection and semantic segmentation. Compared with traditional affine transformation-edge detection, this approach enables joint recognition of water level lines and staff gauge characters, achieving an average recognition error of ± 1.2 cm, with a model accuracy of 93.1%, recall rate of 94.5%, and mAP50:95 of 93.9%. A deep learning-based spectral principal direction recognition method was also employed to calculate the surface water flow velocity, which demonstrated stable and efficient performance, achieving a relative error of 0.005 m/s for the surface velocity. Experimental results confirm that it can effectively address issues such as environmental interference, exhibiting enhanced robustness in low-light and nighttime scenarios. The proposed method provides efficient and accurate identification for dynamic water level monitoring and for real-time detection of river surface flow velocities to improve ecological flow management.

Keywords: ecological flow; water level detection; flow velocity measurement; image recognition; dynamic monitoring

1. Introduction

Ecological flows are key factors in maintaining the health and stability of river ecosystems [1]. The accurate accounting and effective supervision of ecological flows are important for the protection and restoration of ecological systems [2]. Currently, most methods for the dynamic monitoring of river ecological flows rely on contact sensors to provide water level and velocity measurements [3–5]. These devices are not only susceptible to complex horological conditions but also suffer from high maintenance costs and limited measurement accuracy. Moreover, traditional data analysis methods struggle to efficiently extract actionable insights from massive horological datasets, thus failing to meet the demands of real-time ecological flow monitoring. In recent years, the rapid advancement of computer vision, deep learning, and big data technologies has provided innovative approaches to ecological flow monitoring [6–9]. Non-contact measurement methods for water levels and flow rates using image recognition technology can offer advantages such



Academic Editor: Kun Shan

Received: 15 November 2025

Revised: 11 December 2025

Accepted: 22 December 2025

Published: 6 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

as high precision, low cost, and easy maintenance. In addition, deep learning algorithms can uncover hidden patterns in massive datasets, thus significantly improving the accuracy of flow predictions [10–12]. Combined with the Internet of Things and big data cloud-computing technology, real-time monitoring, dynamic analysis, and scientific management of ecological flows can be realized [13–15].

This study employs artificial intelligence algorithms, including deep learning, machine learning, and image recognition, for water level and flow velocity monitoring, which are key parameters of ecological flows. To address the limitations of traditional staff gauge detection methods that combine manual and sensor approaches, a lightweight high-precision model called DSC-YOLOv8n-seg is proposed in this study. The core innovations include replacing the first four conventional convolution modules in the YOLOv8n backbone network with Distributed Shifted Convolution (DSCConv) and Segmentation Head (seg), implementing low-complexity feature extraction through integer operations and offset vectors to reduce the computational load while enhancing the detection of fine features such as staff gauge edges and scale lines, and integrating pixel-level segmentation capabilities into object detection to accurately locate water level boundaries and resolve background interference-induced positioning ambiguities. For challenging river-surface flow velocity measurements, a deep learning-based FFT-STIV (the fast Fourier transform–spatiotemporal image velocimetry) spectral principal direction recognition method based on the DSC-YOLOv8n model is proposed. The proposed model converts velocity measurements into image classification tasks, thus achieving efficient and precise velocity analysis through deep learning integration.

2. Real-Time Water Level Detection Using Image Recognition Technology

2.1. Image Affine Transformation Rotation Correction

Image affine rotation correction is a critical technical challenge in automatic staff gauge image recognition. Its purpose is to correct the image distortion of the staff gauge caused by tilting the camera angle through mathematical transformations, thereby ensuring accurate water level readings [16]. The essence of affine rotation correction lies in establishing a mapping relationship between pixel coordinates and real-world coordinates. By applying linear transformations and translations, tilted staff gauge images can be restored to standard horizontal or vertical perspectives. The specific steps for water level image correction in this study are as follows:

- (1) When extracting the area in which the staff gauge is located from the image, the minimum border box of the labeling tool is used to label and extract the staff gauge to reduce interference from the external environment.
- (2) As the image of the staff gauge in the extraction area is a color (RGB) image, it must be converted into a grayscale image for further analysis. The Canny method is used to automatically select the threshold for binarization and obtain a binary staff gauge image [17].
- (3) Edge detection is used to identify the image edges, followed by the Hough transform for line detection, after which the inclination of the extracted binary image is calibrated. After converting all of the data in the experimental dataset, the average value of the extracted inclination angles is calculated as the inclination angle of the staff gauge in the image [18].
- (4) For the extracted minimum bounding box staff gauge image, the coordinates of the upper-left corner of the bounding box are set as (x_1, y_1) , with a height of h_1 and width of w_1 (Figure 1a). Based on these parameters, the coordinates of the rotation center can be determined as $(x_1 + \frac{w_1}{2}, y_1 + \frac{h_1}{2})$. To obtain the corrected staff gauge image, an affine transformation function is applied to rotate the original image around the

calculated rotation center by θ , the value of which is the averaged tilt angle obtained in the previous step, thereby aligning the staff gauge perpendicular to the water surface (Figure 1b). The rotational transformation principle is described as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_1 - \left(x_1 + \frac{w_1}{2}\right) \\ y_1 - \left(y_1 + \frac{h_1}{2}\right) \end{bmatrix} + \begin{bmatrix} x_1 + \frac{w_1}{2} \\ y_1 + \frac{h_1}{2} \end{bmatrix}, \quad (1)$$

$$M = \text{getRotationMatrix2D}\left(\left(x_1 + \frac{w_1}{2}, y_1 + \frac{h_1}{2}\right), \theta\right), \quad (2)$$

$$\text{img}_{on} = \text{warpAffine}(\text{img}_{in}, M, (w_1, h_1)), \quad (3)$$

where $\begin{bmatrix} x' \\ y' \end{bmatrix}$ is the new coordinate after rotation transformation; $\begin{bmatrix} x_1 + \frac{w_1}{2} \\ y_1 + \frac{h_1}{2} \end{bmatrix}$ is the coordinate of the rotation center; M is the transformation matrix calculated before (rotation matrix); $\text{getRotationMatrix2D}$ is the rotation matrix function; img_{on} and img_{in} are the output and input images, respectively; warpAffine is the affine transformation matrix; and (w_1, h_1) is the size of the output image (width and height of the transformed image).

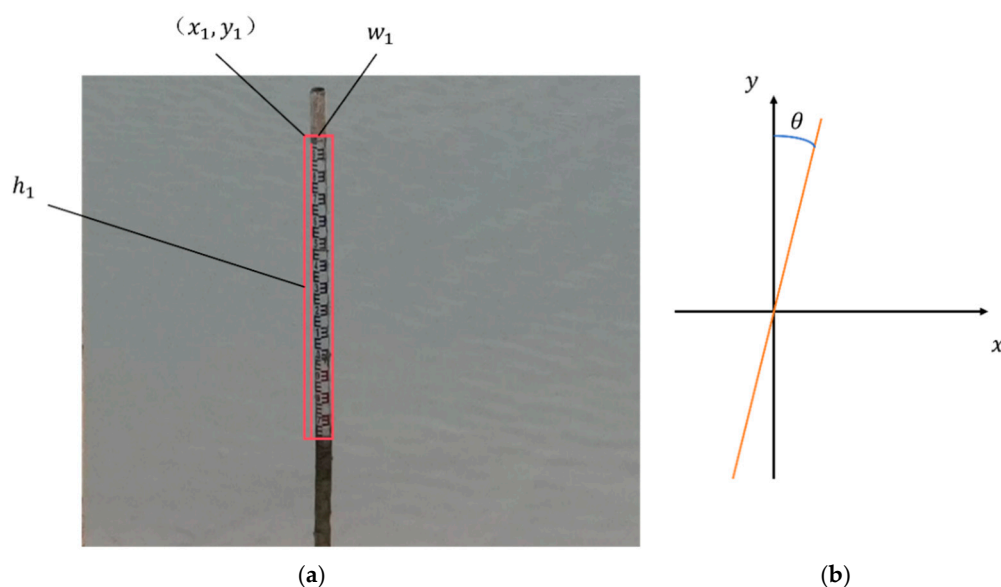


Figure 1. Parameters of the staff gauge tilt detection (a) the coordinates of the upper-left corner of the bounding box (x_1, y_1) , width and height of the transformed image, w_1 and h_1 , and (b) the rotation angle of the original image, θ .

2.2. Grayscale Imaging

When processing staff gauge images, the complex shooting environment presents multiple challenges, including aquatic vegetation obstructions, lighting variations, scale contamination, and water color reflections. To mitigate color inconsistencies across lighting conditions and enhance the contrast between the scale markings, water surface, and background in grayscale space, this study employs a weighted averaging grayscale conversion method. This method effectively reduces the image distortion caused by localized overexposure from intense direct light exposure, thus ensuring clearer staff gauge visuals for subsequent processing. The color images consist of three color channels: red (R), green (G), and blue (B), with each channel typically ranging from 0 to 255. Grayscale is used to abandon the color information and transform the color image into a single-channel gray image; as a result, the staff gauge image only retains the brightness information [19].

The weighted averaging method accounts for the varying contributions of different color channels to visual perception, with the green channel typically having the greatest impact on perception by the human eye. The fundamental principle of this method involves assigning specific weights to each RGB channel and then summing these weights to obtain the grayscale value of the pixel. Figure 2 shows a grayscale image of the staff gauge processed using this method. The specific equation is given as follows:

$$Gray(x, y) = rR(x, y) + gG(x, y) + bB(x, y), \quad (4)$$

where $Gray$, R , G , and B are the pixel values of the grayscale image, red channel, green channel, and blue channel, respectively; the three weight values (r , g , and b) are generally 0.299, 0.587, and 0.114, respectively, based on the differences in the perceptual sensitivities of the three types of cone cells in the human eye's retina to light of different wavelengths.



Figure 2. Grayscale image of the staff gauge after application of the weighted averaging method.

The weighted average method is typically used for standard grayscale processing. The image content and edges of the weighted average method are clearer and more consistent with the perception of grayscale images by the human eye [20]; therefore, it is reasonable to use this method to process the grayscale data of staff gauge images.

2.3. Image Binarization

In water level detection, an appropriate binarization method plays an important role in improving the image detection accuracy [21]. Binarization is a process that simplifies grayscale images by reducing their pixels to two binary values (typically 0 and 255, representing black and white, respectively). The key to this method lies in setting an appropriate threshold [22–24]. By comparing pixel grayscale values with the threshold value, the image can be divided into two distinct regions: the target object (e.g., the water level line or scale markings) and the background. In staff gauge image processing, binarization (or thresholding) converts the pixel values to black or white based on a predefined threshold. This technique is widely used for extracting water level data or segmenting scale markings, making it particularly valuable for hydrological image analyses.

Canny edge detection [25] is an operator that has demonstrated superior performance for measuring water levels in specific environments. Designed to detect edges while minimizing noise interference and achieving precise positioning, it offers rapid response, high accuracy, accurate edge localization, and boundary connectivity capabilities.

In the Gaussian smoothing denoising method, the image is smoothed using a Gaussian filter to reduce the effect of noise on the edge detection results. The Gaussian function is expressed as follows:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right), \quad (5)$$

where σ is the standard deviation of the Gaussian filter, which controls the degree of smoothing.

The gradient intensities and directions are then calculated. The gradient operator is employed to determine the gradient (horizontal gradient E_x , vertical gradient E_y , and total gradient intensity E) and gradient direction of each pixel in the staff gauge image, θ . The relevant equations are as follows:

$$E_x = \sum_{i=-1}^1 \sum_{j=-1}^1 I(i, j) \cdot E_x(i, j), \quad (6)$$

$$E_y = \sum_{i=-1}^1 \sum_{j=-1}^1 I(i, j) \cdot E_y(i, j), \quad (7)$$

$$E = \sqrt{E_x^2 + E_y^2}, \quad (8)$$

$$\theta = \text{atan2}(E_y, E_x), \quad (9)$$

where $I(i, j)$ is the image pixel value; E_x and E_y are the horizontal and vertical gradient operators, respectively; and atan2 is a mathematical function that is widely employed in multiple programming languages to calculate the angle between the line from the origin to the point (x, y) and the positive direction of the x-axis.

The Canny edge detection algorithm employs a dual-threshold mechanism to achieve a refined selection of edge pixels by setting high and low gradient amplitude thresholds. Its core decision criteria can be divided into three levels. When the gradient value of a pixel exceeds the preset high threshold, it is directly identified as a strong edge response and retained. When the gradient value falls between the high and low thresholds, a connectivity analysis is required for verification. Only pixels showing spatial adjacency with confirmed edges are retained; otherwise, suppression is applied. Pixels with gradient values below the low threshold are directly identified as non-edge points and discarded. This hierarchical decision-making strategy effectively balances edge detection sensitivity with noise suppression capability, achieving dual optimization of edge continuity and localization accuracy through threshold coupling mechanisms. Figures 3 and 4 illustrate the threshold-based edge detection workflow and binary image representation, respectively.

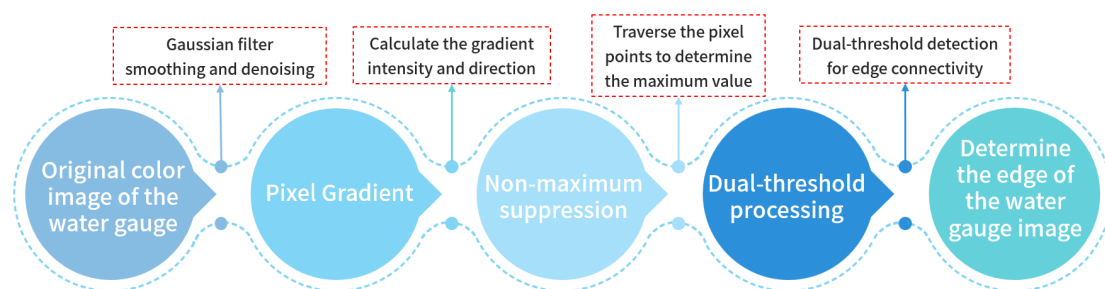


Figure 3. Flow chart of the threshold method for the Canny edge detection algorithm.

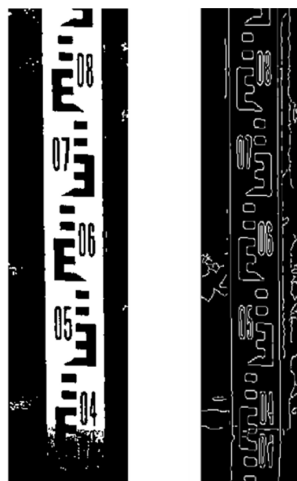


Figure 4. Image of the edge detection threshold binary for the staff gauge.

To address issues such as discontinuous boundaries and white background noise in edge detection and the binarized staff gauge images [26], we employ two complementary morphological operations: erosion and dilation. Although these dual operations are fundamentally opposite in mechanism and essentially serve as inverse processes, both nonlinear algorithms achieve similar image enhancement outcomes. Their shared objective is to effectively eliminate noise and subtle details within the images.

$$E = A \ominus B = \left\{ (x, y) / B_{(x,y)} \subseteq X \right\} \quad E = A \ominus B = \left\{ (x, y) / B_{(x,y)} \subseteq X \right\}, \quad (10)$$

$$D = A \oplus B = \left\{ \frac{(x,y)}{B_{(x,y)}} \cap X \neq \emptyset \right\} \quad D = A \oplus B = \left\{ \frac{(x,y)}{B_{(x,y)}} \cap X \neq \emptyset \right\}, \quad (11)$$

where $\ominus$ and $\oplus$ are the corrosion and expansion operations, respectively; B is the structural unit, A is the corrosion image, and X and D are the connected regions of binary images [27].

2.4. Water Level Determination

The height of the water level gauge in the image will change linearly with the change in water level; therefore, a linear regression model is used to predict the water level value. The steps for predicting the water level value are as follows:

- (1) A training set containing the actual water level gauge height and the corresponding water level values in the images is created. Potential adverse effects caused by individual sample data are eliminated by normalizing the data to a range of 0–1.
- (2) A linear regression model is established.
- (3) The model is trained and tested to obtain the optimal model, and the weight, w , and deviation, b , of the model are obtained.

The linear relationship is defined as follows:

$$y = wx + b, \quad (12)$$

where x is the height of the water level gauge in the image, and y is the water level value.

Linear regression models are typically fitted using the least-squares approximation, which is given as follows:

$$\min_{w,b} E = \frac{1}{N} \sum_{i=1}^N (f(x_i) - y_i)^2, \quad (13)$$

where $\min_{w,b} E$ is the least-squares algorithm.

3. Real-Time Velocity Detection Based on Spatio-Temporal and Spectral Deep Learning Analysis

Another critical parameter sensing technology in ecological flow dynamics monitoring is the real-time detection of the surface flow velocity in rivers. FFT-STIV, a method combining the fast Fourier transform (FFT) and spatiotemporal image velocimetry (STIV) to identify principal directions in spectral images, has been widely applied for river flow monitoring. This study proposes a deep learning-based FFT-STIV spectral image principal direction recognition method that significantly enhances the accuracy and robustness of texture principal direction (MOT) identification.

3.1. Spatio-Temporal Image Generation

STIV is a one-dimensional time-averaged motion vector estimation method with high spatial resolution [28]. Through spatiotemporal correlation analysis, we directly compute the time-averaged velocity vector in specified directions to obtain rapid flow velocity measurements on river surfaces. The process of generating spatiotemporal images is as follows: multiple fixed-length parallel lines aligned with the water flow direction are established on the river video footage as velocity measurement lines, each corresponding to a velocity vector. Continuous video frames (N frames) of river surface images are extracted, and the brightness information from each velocity line is processed to obtain texture-enhanced STI through grayscale variations. Figure 5 illustrates the formation process, where the x-axis represents the velocity line length, and the y-axis denotes the time across N consecutive frames.

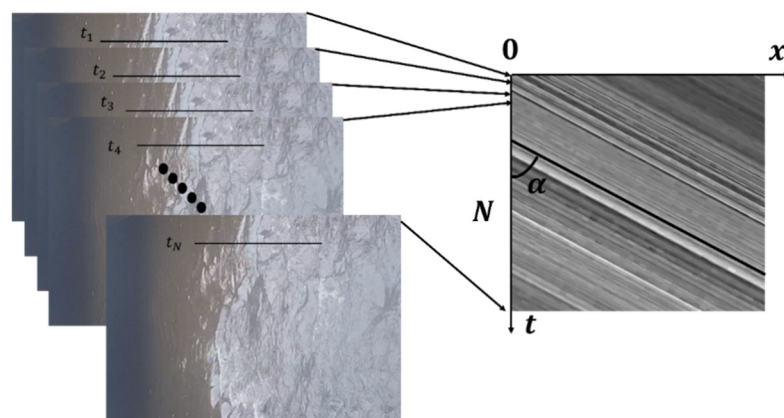


Figure 5. Schematic of spatiotemporal image generation.

3.2. Texture Angle Recognition Method

The FFT-STI method is a spectral analysis technique utilizing FFT to identify the surface flow velocity in spatiotemporal images by analyzing their spectral characteristics [29]. This process involves four key steps. First, spatiotemporal images are generated by extracting specific columns or rows from video sequences. Second, two-dimensional FFT is performed on the generated spatiotemporal images to convert the spatial-domain signals into frequency-domain signals. Third, spectral analysis is employed to identify the dominant frequency components in the spectral images that correspond to textural features and motion information. Fourth, the flow velocity is determined based on these dominant frequency components, and the relationship between the frequency and velocity is mathematically derived using established formulas. Figure 6 shows the spatiotemporal image and frequency image in the STI spatial-to-frequency conversion process.

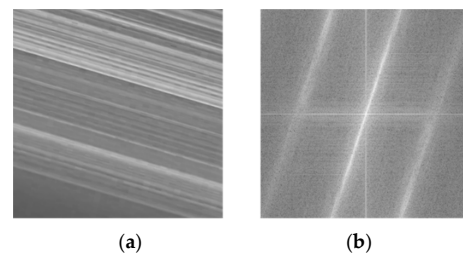


Figure 6. (a) Spatiotemporal image and (b) frequency image in the STI spatial-to-frequency conversion process.

The 2D- $M \times N$ FFT transformation formula for the generated spatiotemporal image is as follows:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{t=0}^{N-1} I(x, t) e^{-j2\pi(ux/M + vt/N)}, \quad (14)$$

$$|F(u, v)| = \left| \sum_{x=0}^{M-1} \sum_{t=0}^{N-1} I(x, t) e^{-j2\pi(ux/M + vt/N)} \right|, \quad (15)$$

where $I(x, t)$ is the value of the image spatial domain; $F(u, v)$ is the spectral image; M and N are the width and height of the spatiotemporal image, respectively; u and v are frequency variables; and $|F(u, v)|$ represents the amplitude of the spectral image [30].

Because the texture angles of the spatiotemporal image generated from the real river data are all distributed within the range of 66° to 84° , the resulting errors are insufficient to verify the generalization and authenticity of the texture angle recognition method. Therefore, artificial spatiotemporal images are selected for validation. The synthesized spatiotemporal images in this study cover angles ranging from 0° to 90° . To achieve accurate detection, 18 different angles are selected at intervals of 5° , with each angle corresponding to one image. Three methods, FFT, QESTA (Quantitative Estimation of Spatiotemporal Texture Anisotropy), and GTM (Gabor Transform Method), are employed to detect the principal texture directions in the simulated spatiotemporal image sequences.

4. Real-Time Water Level and Flow Rate Detection Method

4.1. Hydrological Data Preprocessing

The hydrographic monitoring data comprises two primary sources: observational images from a hydrological monitoring station and field-captured images, totaling 2332 frames. A subset of 2000 images was selected for this study, with 80% allocated to the training set and 20% to the validation set, each containing at least one hydrographic monitoring object. To address challenges such as varying viewing angles and lighting conditions, multiple time periods and camera angles were utilized during field data collection to ensure comprehensive exposure of the hydrographic monitoring features for model training. For object detection tasks, the raw image dataset contained significant variations owing to diverse shooting environments. To mitigate these differences, the Labelling annotation tool was employed to label all of the hydrographic monitoring objects using minimal bounding boxes, with corresponding XML files generated for machine processing.

In this study, a multimodal data preprocessing method is developed for staff gauge images, flow velocity videos, and runoff data. Through image correction, spatiotemporal feature extraction, and data normalization [31], we eliminate redundant information and abnormal interference from the raw data to construct a high-quality dataset. Given that the staff gauge images were collected under varying scenarios, specific preprocessing is required for the images captured under different weather and environmental conditions to ensure high-quality results. For challenging environments, such as low-light conditions

and nighttime settings, histogram equalization is applied to enhance the image quality. Conventional techniques such as binarization and edge detection are employed for normal and enhanced high-quality images. The preprocessing workflow is illustrated in Figure 7.

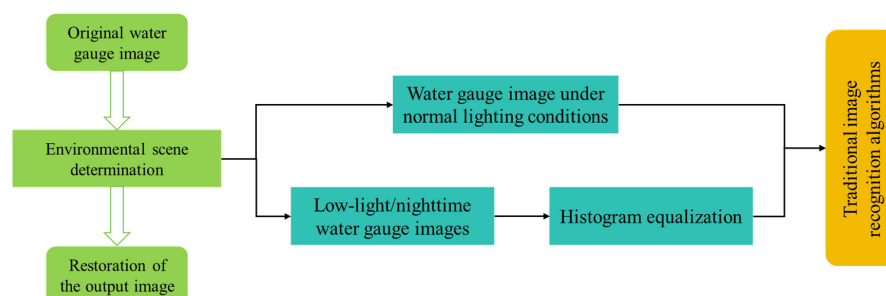


Figure 7. Schematic of the image preprocessing technique.

The core technology involves reconstructing the probability distribution of image pixel values through a grayscale transformation function, thus enabling uniform grayscale levels to effectively improve visual quality. The primary principle is based on the mathematical model used in histogram equalization, the cumulative distribution function (CDF). For the input images, the original histogram is first calculated, followed by its CDF to map the original grayscale levels to the new ones. The processed image and statistical results are shown in Figure 8. The core equation used in the equalization process is

$$s = T\theta = \int_0^r p_r(w)dw, \quad (16)$$

where r is the original grayscale level, $p_r(w)$ is the density function of the original grayscale level, and s is the transformed grayscale level.

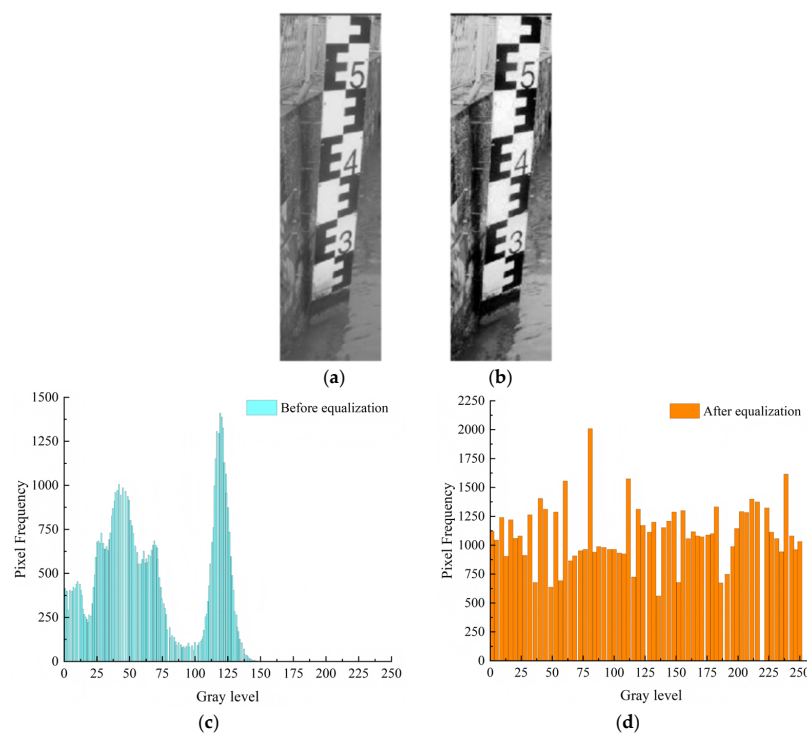


Figure 8. Comparison of grayscale images and histograms before and after histogram equalization. (a) Original grayscale image of the staff gauge; (b) grayscale image of the staff gauge after equalization; (c) grayscale histogram of the staff gauge before equalization and (d) grayscale histogram of the staff gauge after equalization.

4.2. Staff Gauge Detection Evaluation Indices

In staff gauge image detection, commonly used evaluation metrics include the precision (P), recall (R), and mean average precision (mAP). Precision measures the true detection rate of staff gauge recognition models, whereas recall assesses the false detection rates. Higher P and R values indicate better measurement accuracy.

Additionally, the intersection–union ratio (IoU) measures the overlap between the predicted and true bounding boxes, with values ranging from zero to one [32].

4.3. Design of Flow Rate Detection Experiments and Evaluation Indices

Continuous video data from January 2019 to February 2020, covering different weather conditions (rainy days and sunny days) and different time periods (excluding nighttime values), were used. To validate the feasibility of the proposed scheme, we conducted comparative verification using a series of collected river videos. The acquired footage featured eight velocity measurement lines captured by equipment with a resolution of 1920 (1080p, 20 FPS) in each 30 s segment. Frame-by-frame image extraction yielded a total of 600 image sequences. To demonstrate the detection accuracy of the model, we randomly selected two datasets: one from River C (16 March 2019, 10:15) and the other from River S (6 October 2020, 15:15) (single-frame screenshots of both rivers are shown in Figure 9). Each dataset contained 600 spectral images, totaling 1200 images for testing and validation. These images contained eight velocity measurement lines, yielding 9600 spatial-frequency images when categorized by measurement lines. The average velocity of the eight velocity lines was calculated to determine the final surface flow velocity in the measurement section.



Figure 9. Single-frame screenshots of two rivers during the test period: (a) River S and (b) River C.

When evaluating the performance of the model, although this study converts the main direction recognition of the spectral image into a classification task and takes the final prediction result as the recognition angle, the evaluation indices of the regression prediction model (the mean absolute error (MAE), mean square error (MSE), and mean relative error (MRE)) can still be used to comprehensively and accurately evaluate the performance of the model [33]. The hyperparameter settings for each model in the model training stage are listed in Table 1.

Table 1. Hyperparameters of each model.

Model	Optimizer	Loss Function	Batch Size	Epoch	Learning Rate
ResNet18	SGD	cross entropy	32	300	0.001
SSD	SGD	cross entropy	32	300	0.001
YOLOv5m	Adam	cross entropy	32	300	0.001
DSC-YOLOv8n	SGD	cross entropy	32	300	0.01

Notes: The “Optimizer” refers to the algorithms used during model training, such as SGD (Stochastic Gradient Descent) and Adam. The choice of optimizer affects the model’s convergence speed and final performance. The “Loss function” is a function used to measure the difference between the model’s predicted values and the true values. The “Batch Size” represents the number of samples used in one iteration of training. The “Epoch” represents the number of times the entire training dataset is fully used. The “Learning rate” is a key hyperparameter that controls the step size of model parameter updates.

5. Comparison of Test Results and Performance Analysis

5.1. Staff Gauge Detection Experiment

This study analyzed the experimental results obtained by replacing regular blocks with DSConv in the backbone network of YOLOv8n and identified the optimal replacement configuration as DSC-YOLOv8n. Comparative evaluations of the training cycles of DSC-YOLOv8n with those of both the base YOLOv8n model and other classic YOLO family models were conducted to assess the performance. Visualizations of the prediction processes and detection results in complex scenarios for both DSC-YOLOv8n and the base YOLOv8n were presented, demonstrating the robustness and effectiveness of the model in water-scale image detection.

(1) Fusion experiment

To identify the optimal improved model, we conducted a series of ablation experiments based on YOLOv8n. The results are presented in Table 2. Specifically, we replaced all of the Conv modules in the YOLOv8n-seg backbone network with DSConv variants and analyzed the contribution of each substitution to the detection performance enhancement. The table below provides a comparative analysis of all the replacement options. The selected Conv modules were replaced with DSConv, whereas the others were replaced with the original modules.

As the Conv modules in the YOLOv8n-seg backbone network were replaced by DSConv, the minimum mAP50:95 value reached 86.7% when the fifth Conv module was substituted. When the first, second, and fifth Conv modules were replaced with DSConv, the mAP50 value dropped to a minimum of 87.5%. However, when the first, second, third, and fourth Conv modules were replaced with DSConv, the mAP50 value surged to a peak of 93.1%, with the maximum mAP50:95 reaching 93.9%. Consequently, the substitution of the first four Conv modules with DSConv was recognized as the optimal improvement outcome, resulting in the DSC-YOLOv8n-seg configuration.

(2) Comparison between the basic YOLOv8n model and DSC-YOLOv8n-seg

To evaluate the performance of DSC-YOLOv8n-seg, a comparative analysis was conducted with the base YOLOv8n model using identical datasets and experimental configurations. Table 3 presents the performance comparison between YOLOv8n and DSC-YOLOv8n-seg. The results demonstrate that DSC-YOLOv8n-seg significantly outperformed the base YOLOv8n model, achieving a 4.2% improvement in mAP50 and a 6.2% enhancement in mAP50:95.

Table 2. Ablation Experiment Results of Different DSConv Module Replacement Schemes.

Conv-1	Conv-2	Conv-3	Conv-4	Conv-5	mAP50	mAP50:95
					88.9	87.7
✓					89.5	88.1
	✓				90.0	88.8
		✓			90.5	89.5
			✓		89.0	87.4
				✓	90.2	86.7 ↓
✓	✓				91.2	86.0
✓		✓			90.7	85.3
✓			✓		91.7	84.7
✓				✓	91.5	85.3
	✓	✓			91.0	85.9
	✓		✓		90.5	86.5
	✓			✓	89.5	87.0
		✓	✓		90.0	87.6
		✓		✓	89.5	88.2
			✓	✓	89.0	88.7
✓	✓	✓			88.5	89.2
✓	✓		✓		88.0	89.7
✓	✓			✓	87.5 ↓	90.2
✓		✓	✓		88.3	90.6
✓		✓		✓	89.1	91.1
✓			✓	✓	90.2	91.8
	✓	✓	✓		90.7	92.6
	✓	✓		✓	91.2	92.4
	✓		✓	✓	91.7	91.6
		✓	✓	✓	92.6	92.1
✓	✓	✓	✓		93.1 ↑	93.9 ↑
✓	✓	✓		✓	92.1	93.1
✓	✓		✓	✓	91.5	92.8
✓		✓	✓	✓	90.5	91.9
	✓	✓	✓	✓	89.6	91.3
✓	✓	✓	✓	✓	88.8	89.9

Notes: “✓” indicates that the corresponding Conv module is replaced. Numbers in red font indicate that the accuracy of the model under the corresponding alternative scheme has experienced a relatively significant decrease or increase. Among them, a downward arrow denotes a decline in accuracy, while an upward arrow denotes an improvement in accuracy. A red check mark indicates the alternative scheme selected as the final option. Bold formatting indicates the scheme shows the best performance.

Table 3. Performance Comparison between the Improved Model DSC-YOLOv8n-seg and the Baseline Model YOLOv8n.

Model	mAP50 (%)	mAP50:95 (%)
YOLOv8n	0.889	0.877
DSC-YOLOv8n-seg	0.931	0.939

Figure 10 shows a detailed comparison between the training processes of DSC-YOLOv8n-seg and the baseline YOLOv8n model, with particular emphasis on four key metrics: accuracy, recall rate, mAP50, and mAP50:95. The analysis reveals that as the number of epochs increased, both DSC-YOLOv8n-seg and YOLOv8n ultimately reached a state of convergence during their training processes.

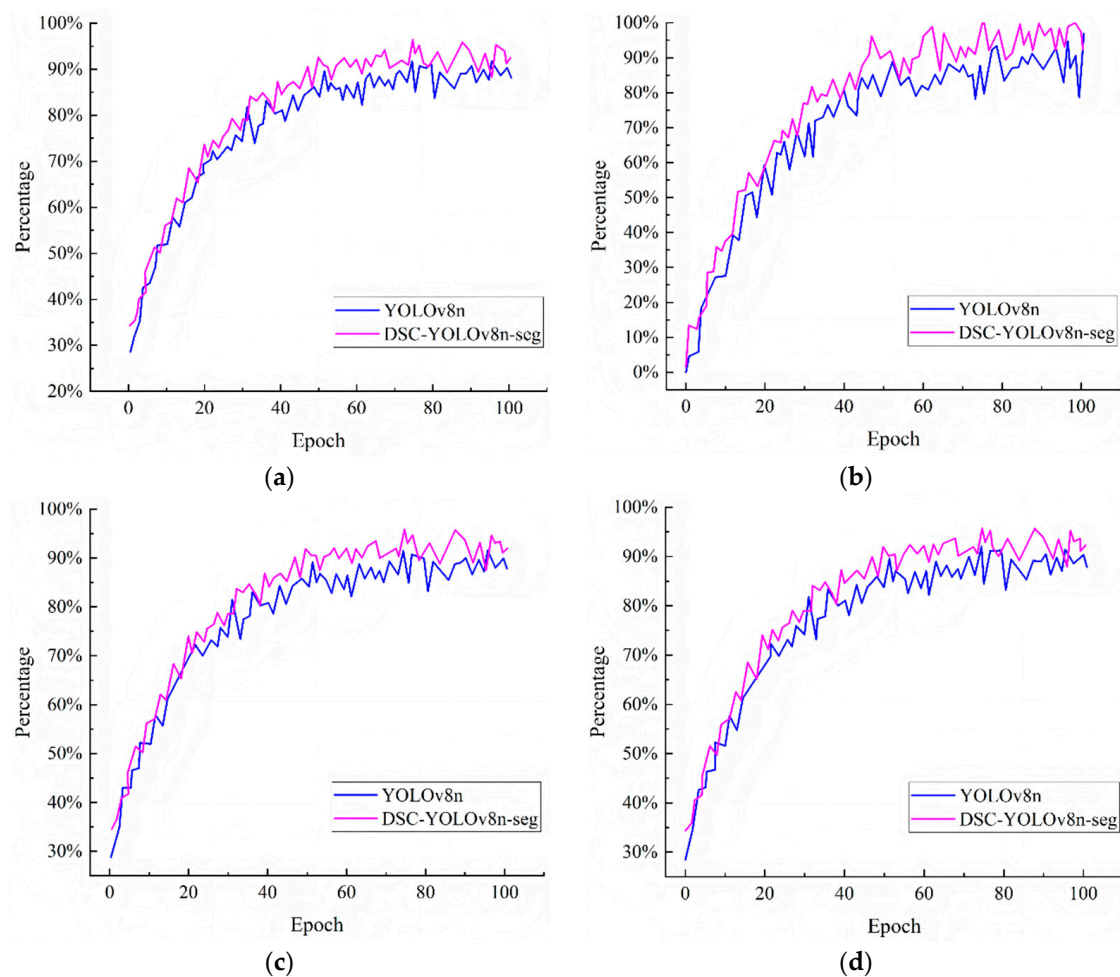


Figure 10. Comparison of (a) Precision, (b) Recall Percentage, (c) MAP50 and (d) MAP50:95 during model training.

Specifically, Figure 10a illustrates the training process of the precision index. The DSC-YOLOv8n-seg and YOLOv8n models alternated in the first 30 epochs, after which DSC-YOLOv8n-seg demonstrated significantly better accuracy. Figure 10b shows that the recall index of the DSC-YOLOv8n-seg model consistently outperformed that of the base model. Figure 10c,d clearly demonstrate that the mAP50 and mAP50:95 exhibited similar training curves. During the first 40 epochs, the performance of DSC-YOLOv8n and the base YOLOv8n model remained closely matched, after which DSC-YOLOv8n demonstrated a notable improvement.

In conclusion, the evaluation of these metrics for both the baseline YOLOv8n and DSC-YOLOv8n-seg models demonstrates that the improved model achieves higher performance, thereby validating the effectiveness of the enhancements.

(3) Comparison between DSC-YOLOv8n-seg and other models

Using the same staff gauge data, the DSC-YOLOv8n-seg model, the traditional deep convolutional network model, and other YOLO target detection models were compared to further verify the performance advantages of the proposed staff gauge image detection method.

This study investigated staff gauge level detection across different methodologies. As presented in Table 4, we first compared DSC-YOLOv8n-seg with traditional deep convolutional networks and other YOLO detection models (including YOLOv5m, YOLOv8n-seg, and DSC-YOLOv8n). The results demonstrate that DSC-YOLOv8n-seg exhibited significant

performance improvements. Regarding measurement errors, four types of deviations were analyzed. The results reveal that conventional deep convolutional networks exhibited relatively large errors, averaging approximately 2.6 cm. In contrast, the improved DSC-YOLOv8n-seg model achieved errors of less than 1 cm for 56% of the measurements, errors of 1–3 cm for 33% of the measurements, and errors exceeding 3 cm in only 11% of the measurements; the average error was 1.2 cm. These metrics confirm the superior capability of DSC-YOLOv8n-seg for staff gauge level recognition.

Table 4. Comparison of Water Level Recognition Error Distribution and Average Error of Different Detection Models.

Model	Error $\leq$ 1 cm	1 cm < Error < 3 cm	Error > 3 cm	Average Error (cm)
Deep convolutional networks	14%	49%	37%	2.6
YOLOv5m	21%	57%	22%	2.2
YOLOv8n	21%	51%	28%	2.3
YOLOv8n-seg	25%	56%	19%	1.9
DSC-YOLOv8n	46%	37%	17%	1.7
DSC-YOLOv8n-seg	56%	33%	11%	1.2

A performance analysis of the models is presented in Table 5. Specifically, the DSC-YOLOv8n-seg model achieved a precision (P) of 93.1%, recall (R) of 94.5%, mean average precision at 50 (mAP50) of 93.1%, and mAP50:95 of 93.9%, achieving the best performance among the compared models.

Table 5. Comparison of Model Performance of the DSC-YOLOv8n-seg Model and other models.

Model	Accuracy (%)	Recall (%)	mAP50 (%)	mAP50:95 (%)
Deep convolutional networks	85.1	86.7	84.9	85.3
YOLOv5m	87.8	87.9	87.5	88.2
YOLOv8n	88.2	93.7	88.9	87.7
YOLOv8n-seg	89.4	90.5	91.3	90.6
DSC-YOLOv8n	91.2	92.3	91.6	92.7
DSC-YOLOv8n-seg	93.1	94.5	93.1	93.9

To evaluate the staff gauge recognition across three scenarios—daytime, nighttime, and low-light conditions—this study employed three distinct environments. During daytime operation, the model demonstrated optimal performance with average errors of approximately 0.8 cm, owing to sufficient illumination and clear staff gauge markings. Under nighttime conditions, the enhanced model maintained accurate detection using auxiliary light sources such as equipment illumination and night lights. In low-light scenarios, the system achieved reliable recognition despite slightly dimmed images, maintaining sufficient clarity for effective detection. The recognition accuracy of the model under these conditions is presented in Table 6.

Table 6. Recognition accuracy of DSC-YOLOv8n-seg in different lighting scenarios.

Scene	Error $\leq$ 1 cm	1 cm < Error < 3 cm	Error $\geq$ 3 cm	Average Error (cm)
Daylight	88%	11%	1%	0.8
Low light	40%	28%	32%	1.9
Night	58%	42%	0%	0.9

5.2. Flow Rate Detection Experiment

The results are shown in Tables 7 and 8, which compare the river velocities calculated and observed.

Table 7. Comparison of various velocity measurement lines on River C recorded on 16 March 2019, 10:15.

Speedometer Line	MOT	V_{img}	$V_{img-avg}$	V_{truth}	Relative Error
1	72	0.428	0.433	0.429	0.4%
2	72	0.431			
3	72	0.436			
4	71	0.422			
5	71	0.429			
6	72	0.443			
7	71	0.439			
8	73	0.437			

Table 8. Comparison of various velocity measurement lines on River S recorded on 6 October 2020, 15:15.

Speedometer Line	MOT	V_{img}	$V_{img-avg}$	V_{truth}	Relative Error
1	83	0.186	0.185	0.190	0.5%
2	82	0.197			
3	84	0.174			
4	81	0.173			
5	83	0.194			
6	83	0.183			
7	82	0.182			
8	83	0.191			

Note: MOT is the textural angle, V_{img} indicates the measurement result of the surface flow velocity identified by DSC-YOLOv8n, $V_{img-avg}$ is the average flow velocity, and V_{truth} is the actual flow velocity.

Through a comparative verification of randomly selected data from both groups, the two tables above demonstrate that DSC-YOLOv8n ultimately calculated the average river flow velocities with relative errors of 0.004 and 0.005 m/s compared to the actual surface flow velocities. This result confirms the effectiveness of DSC-YOLOv8n for accurately calculating the average river surface flow velocities during spectral image principal direction recognition using FFT-STIV technology, while maintaining high precision.

6. Conclusions

The proposed staff gauge level and river flow velocity image recognition method enables real-time acquisition of hydrological data. Ecological flow monitoring requires the dynamic sensing of critical parameters such as water levels and velocities. Traditional contact-based measurement methods suffer from low accuracy and maintenance challenges in complex hydrological environments; thus, these methods cannot meet the real-time supervision demands of the Yellow River Basin. This study integrates computer vision and deep learning technologies to develop a non-contact multi-parameter monitoring system: a lightweight segmentation model, DSC-YOLOv8n-seg, is constructed for staff gauge detection to achieve pixel-level water level line localization. For surface flow velocity measurement, a spectral principal direction recognition method based on DSC-YOLOv8n is designed using FFT-STIV, transforming the velocity measurement into an image classification problem. Through ablation experiments and multi-scenario validations, our approach is shown to significantly enhance the measurement accuracy and environmental adaptability, thus providing high-precision “dual hydrodynamic parameter” inputs

for ecological flow accounting and driving the intelligent transformation of traditional hydrological monitoring systems.

Author Contributions: Conceptualization, J.Y. and Y.L.; methodology, T.W. and P.Z.; experiment and performance analysis, W.J. and L.X.; resources, Y.L.; data curation, T.W.; writing—original draft preparation, Y.L. and P.Z.; writing—review and editing, J.Y.; visualization, P.Z.; supervision, J.Y. and P.Z.; funding acquisition, J.Y. and P.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded by the National Key Research and Development Program of China (2023YFC3209200); and the Fundamental Research Funds for the Central Universities (B240201103).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: Authors Jun Yu, Wenlong Jiang, and Lei Xing were employed by the company Yellow River Engineering Consulting Co., Ltd. Authors Yongsheng Li and Ting Wang were employed by the company Yunhe (Henan) Information Technology Co., Ltd. The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

1. Liu, S.; Zhang, Q.; Xie, Y.; Xu, P.; Du, H. Evaluation of Minimum and Suitable Ecological Flows of an Inland Basin in China Considering Hydrological Variation. *Water* **2023**, *15*, 649. [\[CrossRef\]](#)
2. Mo, C.; Wang, H.; Lai, S.; Huang, Y.; Gao, W.; Tang, L.; Zhong, S. Early Warning Forecasting of River Ecological Flow in a Typical Karst Basin, Southwest China: Leveraging Baseflow Segmentation and Interpretable Machine Learning. *J. Hydrol. Reg. Stud.* **2025**, *59*, 102433. [\[CrossRef\]](#)
3. Zhang, Q.; Zhang, Z.; Shi, P.; Singh, V.P.; Gu, X. Evaluation of Ecological Instream Flow Considering Hydrological Alterations in the Yellow River Basin, China. *Glob. Planet. Change* **2017**, *160*, 61–74. [\[CrossRef\]](#)
4. Ni, X.; Dong, Z.; Xie, W.; Wu, S.; Chen, M.; Yao, H.; Jia, W. A Practical Approach for Environmental Flow Calculation to Support Ecosystem Management in Wujiang River, China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 11615. [\[CrossRef\]](#)
5. Wynants, M.; Hallberg, L.; Prischl, L.A.; Livsey, J.; Bieroz, M. Trends and Purposes of European River Monitoring and Restoration. *Environ. Sci. Policy* **2025**, *170*, 104130. [\[CrossRef\]](#)
6. Sahoo, D.P.; Sahoo, B.; Tiwari, M.K.; Behera, G.K. Integrated Remote Sensing and Machine Learning Tools for Estimating Ecological Flow Regimes in Tropical River Reaches. *J. Environ. Manag.* **2022**, *322*, 116121. [\[CrossRef\]](#)
7. Fang, C.; Yuan, G.; Zheng, Z.; Zhong, Q.; Duan, K. Technical Note: Monitoring Discharge of Mountain Streams by Retrieving Image Features with Deep Learning. *Hydrol. Earth Syst. Sci.* **2024**, *28*, 4085–4098. [\[CrossRef\]](#)
8. Porras, A.A.; Jin, Y.H.; Peacock, E.; Avila, Y.; Chu, A.N.; Mahaseth, H.; Santiago, A.L.; Siegmund, J.G. Expanding an Urban Stream Ecological Monitoring System with Functional Data and Deep Learning. *Freshw. Sci.* **2024**, *43*, 189–205. [\[CrossRef\]](#)
9. Li, M.; Zhao, C.; Huang, Q.; Pan, T.; Yesou, H.; Nerry, F.; Li, Z.L. Combining Landsat 5 TM and UAV Images to Estimate River Discharge with Limited Ground-Based Flow Velocity and Water Level Observations. *Remote Sens. Environ.* **2025**, *318*, 114610. [\[CrossRef\]](#)
10. Campos Filho, L.C.P.; Figueiredo, N.M.d.; Blanco, C.J.C.; Tobias, M.S.G.; Afonso, P. Machine Learning for the Sustainable Management of Depth Prediction and Load Optimization in River Convoys: An Amazon Basin Case Study. *Sustainability* **2024**, *16*, 8517. [\[CrossRef\]](#)
11. Fu, B.; Liu, Y.; Li, Y.; Wang, C.; Li, C.; Jiang, W.; Hua, T.; Zhao, W. The Research Priorities of Resources and Environmental Sciences. *Geogr. Sustain.* **2021**, *1*, 87–94. [\[CrossRef\]](#)
12. Wang, X.; Li, Z.; Zhang, Y.; An, G. Water Level Recognition Based on Deep Learning and Character Interpolation Strategy for Stained Water Gauge. *River* **2023**, *2*, 506–517. [\[CrossRef\]](#)
13. Manfreda, S.; Miglino, D.; Saddi, K.C.; Jomaa, S.; Eltner, A.; Perks, M.; Peña-Haro, S.; Bogaard, T.; van Emmerik, T.H.; Mariani, S.; et al. Advancing River Monitoring Using Image-Based Techniques: Challenges and Opportunities. *Hydrol. Sci. J.* **2024**, *69*, 657–677. [\[CrossRef\]](#)
14. Ding, Z.; Zheng, K.; Zhao, R.; Mei, J.; Hou, L. Design and Implementation of an IoT System for Monitoring Ecological Environment. *IEEE Access* **2024**, *12*, 141324–141334. [\[CrossRef\]](#)
15. Huang, Y.; Chen, H.; Liu, B.; Huang, K.; Wu, Z.; Yan, K. Radar Technology for River Flow Monitoring: Assessment of the Current Status and Future Challenges. *Water* **2023**, *15*, 1904. [\[CrossRef\]](#)

16. Kim, J.D.; Han, Y.J.; Hahn, H.S. Image-Based Water Level Measurement Method under Stained Ruler. *J. Meas. Sci. Instrum.* **2010**, *1*, 28–31.
17. Tawfeeq, N.; Harbi, J. Using the Canny Method with Deep Learning for Detect and Predict River Water Level. *J. Al-Qadisiyah Comput. Sci. Math.* **2024**, *16*, 135–150. [[CrossRef](#)]
18. Zhao, K.; Han, Q.; Zhang, C.-B.; Xu, J.; Cheng, M.-M. Deep Hough Transform for Semantic Line Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 4793–4806. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, W.; Jin, S.; Zhuang, P.; Liang, Z.; Li, C. Underwater Image Enhancement via Piecewise Color Correction and Dual Prior Optimized Contrast Enhancement. *IEEE Signal Process. Lett.* **2023**, *30*, 229–233. [[CrossRef](#)]
20. Li, Z.; Yin, C.; Zhang, X. Crack Segmentation Extraction and Parameter Calculation of Asphalt Pavement Based on Image Processing. *Sensors* **2023**, *23*, 9161. [[CrossRef](#)]
21. Jiang, X.; Tong, Z.; Yu, Z.; Jiang, P.; Xu, L.; Wu, L.; Chen, M.; Zhang, Y.; Zhang, J.; Yang, X. Fourier Single-Pixel Imaging Based on Online Modulation Pattern Binarization. *Photonics* **2023**, *10*, 963. [[CrossRef](#)]
22. Rezanezhad, V.; Baierer, K.; Neudecker, C. A Hybrid CNN-Transformer Model for Historical Document Image Binarization. In Proceedings of the 7th International Workshop on Historical Document Imaging and Processing, New York, NY, USA, 25–26 August 2023.
23. Michalak, H.; Krupiński, R.; Lech, P.; Okarma, K. Preprocessing of Document Images Based on the GGD and GMM for Binarization of Degraded Ancient Papyri Images. In *Progress in Image Processing, Pattern Recognition and Communication Systems*; Springer: Cham, Switzerland, 2022.
24. Kang, K.U.S. Complex Image Processing with Less Data—Document Image Binarization by Integrating Multiple Pre-Trained U-Net Modules. *Pattern Recognit.* **2021**, *109*, 107577. [[CrossRef](#)]
25. Song, Y.; Li, C.; Xiao, S.; Zhou, Q.; Xiao, H. A Parallel Canny Edge Detection Algorithm Based on OpenCL Acceleration. *PLoS ONE* **2024**, *19*, 31. [[CrossRef](#)] [[PubMed](#)]
26. Jin, X.C.; Ong, S.H.; Jayasooriah, J. A Domain Operator for Binary Morphological Processing. *IEEE Trans. Image Process.* **1995**, *4*, 1042–1046. [[CrossRef](#)] [[PubMed](#)]
27. Li, J.; Tong, C.; Yuan, H.; Huang, W. A Complex Environmental Water-Level Detection Method Based on Improved YOLOv5m. *Sensors* **2024**, *24*, 5235. [[CrossRef](#)]
28. Zhao, H.; Chen, H.; Liu, B.; Liu, W.; Xu, C.-Y.; Guo, S.; Wang, J. An Improvement of the Space-Time Image Velocimetry Combined with a New Denoising Method for Estimating River Discharge. *Flow Meas. Instrum.* **2021**, *77*, 101864. [[CrossRef](#)]
29. Perks, M.T.; Sasso, D.; Hauet, A.; Coz, J.L.; Manfreda, S. Towards Harmonization of Image Velocimetry Techniques for River Surface Velocity Observations. *Earth Syst. Sci. Data* **2020**, *12*, 1545–1559. [[CrossRef](#)]
30. Chen, M.; Chen, H.; Wu, Z.; Huang, Y.; Zhou, N.; Xu, C.Y. A Review on the Video-Based River Discharge Measurement Technique. *Sensors* **2024**, *24*, 4655. [[CrossRef](#)]
31. Zhang, S.J.; Zhu, K.; Sun, X.Y.; Li, D.S. The role of matching pursuit algorithm and multi-scale daily rainfall data obtained from decomposition in runoff prediction. *J. Hydrol.* **2024**, *53*, 101836. [[CrossRef](#)]
32. Lv, X.; Guo, H.; Tian, Y.; Meng, X.; Bao, A.; De Maeyer, P. Evaluation of GSMaP Version 8 Precipitation Products on an Hourly Timescale over Mainland China. *Remote Sens.* **2024**, *16*, 210. [[CrossRef](#)]
33. Yang, Y.; Xu, L.; Luo, M.; Wang, X.; Cao, M. Detection Algorithm of Laboratory Personnel Irregularities Based on Improved YOLOv7. *Comput. Mater. Contin.* **2024**, *78*, 2741–2765. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.