Article

Super-Resolution Semantic Segmentation of Droplet Deposition Image for Low-Cost Spraying Measurement

Jian Liu [1], **Shihui Yu** [1], **Xuemei Liu** [1,2,*], **Guohang Lu** [1], **Zhenbo Xin** [1,2] **and Jin Yuan** [1,2]

1   College of Mechanical & Electronic Engineering, Shandong Agricultural University, Tai'an 271018, China; unknownlogic@163.com (J.L.); yushihuilearn@163.com (S.Y.); gh.lu_701@sdau.edu.cn (G.L.); xinzb@sdau.edu.cn (Z.X.); jyuan@sdau.edu.cn (J.Y.)
2   Shandong Agricultural Equipment Intelligent Engineering Laboratory, Tai'an 271018, China
*   Correspondence: lxmywj@126.com; Tel.: +86-152-6986-8559

Abstract: In-field in situ droplet deposition digitization is beneficial for obtaining feedback on spraying performance and precise spray control, the cost-effectiveness of the measurement system is crucial to its scalable application. However, the limitations of camera performance in low-cost imaging systems, coupled with dense spray droplets and a complex imaging environment, result in blurred and low-resolution images of the deposited droplets, which creates challenges in obtaining accurate measurements. This paper proposes a Droplet Super-Resolution Semantic Segmentation (DSRSS) model and a Multi-Adhesion Concave Segmentation (MACS) algorithm to address the accurate segmentation problem in low-quality droplet deposition images, and achieve a precise and efficient multi-parameter measurement of droplet deposition. Firstly, a droplet deposition image dataset (DDID) is constructed by capturing high-definition droplet images and using image reconstruction methods. Then, a lightweight DSRSS model combined with anti-blurring and super-resolution semantic segmentation is proposed to achieve semantic segmentation of deposited droplets and super-resolution reconstruction of segmentation masks. The weighted IoU (WIoU) loss function is used to improve the segmented independence of droplets, and a comprehensive evaluation criterion containing six sub-items is used for parameter optimization. Finally, the MACS algorithm continues to segment the remained adhesive droplets processed by the DSRSS model and corrects the bias of the individual droplet regions by regression. The experiments show that when the two weight parameters α and β in WIoU are 0.775 and 0.225, respectively, the droplet segmentation independence rate of DSRSS on the DDID reaches 0.998, and the IoU reaches 0.973. The MACS algorithm reduces the droplet adhesion rate in images with a coverage rate of more than 30% by 15.7%, and the correction function reduces the coverage error of model segmentation by 3.54%. The parameters of the DSRSS model are less than 1 M, making it possible to run it on embedded platforms. The proposed approach improves the accuracy of spray measurement using low-quality droplet deposition image and will help to scale-up of fast spray measurements in the field.

Keywords: droplet deposition measurement; image processing; super-resolution; semantic segmentation

1. Introduction

Precision agriculture, arising from the need to mitigate the environmental impact and enhance agricultural input efficiency, necessitates the meticulous control of pesticide application. This places significant demands on the spraying quality of land sprayers and drones. In order to carry out high-quality spraying operations, on the one hand, new types of sprayers with higher performance have been developed [1], and on the other hand, it is necessary to monitor the quality of the spraying in real time, in particular to monitor droplet deposition on the leaves of the plant in a comprehensive manner. The monitoring measurements can then be utilized as feedback data to support the adjustment of the sprayer's spraying parameters accordingly, which serves as the key to reducing pesticide

application and improving overall efficiency. To ensure the real-time and convenient monitoring of plant droplet deposition, it becomes essential to enhance the automatic digitization of measuring droplet deposition on plant leaves. Furthermore, it is crucial that we reduce the time delay associated with individual measurements and develop rapid in situ measurement techniques. These endeavors have emerged as a focal point of research in this field [2–6].

Indicators for measuring spray quality include droplet coverage, droplet density, Volume Median Diameter (VMD), Number Median Diameter (NMD), etc. Calculating these indicators requires counting the number of droplets and the volume or diameter of each individual droplet. However, in recent years, as drones and sprayers have widely used precise nozzles for high-quality pesticide application operations, the droplets from agricultural sprayers are generally relatively dense, and the evaluation needs to be completed with the help of sophisticated test strips or devices. Currently, the spray measurement method with the widest range of applications and the most mature technology is the water-sensitive paper method (WSP). The process of implementing this method includes steps such as placement, spraying, recycling, scanning, and analyzing. In the early application of the WSP method, scanning and analyzing needed to be performed in the laboratory, and transferring the water-sensitive paper consumed a lot of time. Zhu et al. used a portable business card scanner, a portable computer, and a customized software package called DepositScan version 1.2 to form a system which digitized the method to a certain extent and greatly improved the speed [7]. However, the above method still relies on the manual operation of WSP, such as placing and retrieving and discrete digitalization, and the consumption of WSP still restricts the sustainability of this method. The consumption of water-sensitive paper during large-scale droplet measurement brings excessive economic costs.

Recently, machine vision has been utilized in spray measurement to acquire the details of droplets deposited by agricultural sprays. Wang et al. designed a new piece of machine vision-based droplet collection equipment, using recyclable and erasable oil paper instead of WSP [4]. The device is less disturbed by the environment and can realize automatic control and the automation of spray deposition image shooting. The collected images are segmented by a watershed algorithm to achieve droplet segmentation. Based on the above equipment, Wang, et al. proposed a droplet detection model based on SSD_MobileNet, which reconstructed the droplet outline based on the detection results, further improving the statistical accuracy of the droplet parameters [8]. Yang, et al. used the deep learning semantic segmentation model Deeplab V3+ to realize the segmentation of deposited droplets, ensuring the accuracy of droplet segmentation [6]. This study considers clusters of adherent droplets and independent droplets as two different object categories. The neural network is used to find the concave points in the adherent droplets, and the concave segmentation algorithm is used to individually separate the adherent droplet clusters. The rate of recognizing adhesion pits reaches 95%, which can effectively identify and segment most of the adhered droplets, thus effectively improving the accuracy of the spray measurement.

When droplets are deposited on WSP and erasable oil paper, they are in an absorption and diffusion state, which is completely different from the form of droplets deposited on leaves. The area which is stained by the deposition of spray on WSP and erasable oil paper is larger than the area covered on the leaf surface. Under a high-intensity spray, it is easy for the overlap of stained patches to occur, affecting the measurement results. Using a non-water-absorbent material as the surface for the deposition of the droplets can simulate the hemispherical state of droplets when deposited on leaves. Vacalebre, et al. studied the wetting characteristics of acrylic material surfaces. It was found that the acrylic surface treated with a specific method can obtain strong hydrophobicity and a stable WCA (wetting angle) [9]. The volume of a droplet can be calculated from its deposition area using the known wetting agent angle [10]. According to the above principle, an Optical Droplet Edge Imaging (ODEI)-based automatic spray measurement method has been designed. Its measurement accuracy is directly related to the performance of the

camera equipped with the measurement device, but high-performance cameras are not conducive to their promotion in scalable application. However, the limitations caused by camera performance in a low-cost imaging system, coupled with dense spray droplets and a complex imaging environment, result in blurred and low-resolution images of deposited droplets, which makes obtaining accurate measurements challenging. How to achieve precise measurements by leveraging low-cost cameras becomes a practical research topic.

In recent years, deep learning has made great progress in the fields of image super-resolution reconstruction and deblurring. Image deblurring models such as DeblurGAN [11,12] can repair image blur, and its effect is better than traditional methods such as Wiener filtering and deconvolution; Quan, et al. proposed using a linear mixture of a series of Gaussian blur kernels to approximately express out-of-focus blur, and built the end-to-end deep learning model GKMNet to correct de-focus blur with remarkable results [13]. Super-resolution reconstruction models such as SRGAN [14] can reconstruct image details when enlarging the image, and its effect is far better than traditional methods based on interpolation. The deep learning segmentation model can consider the semantic information of pixels when segmenting images and identify light spots as being parts of droplets, which is better than traditional methods. Some research combines super-resolution ideas with semantic segmentation methods to improve segmentation accuracy [15].

This paper reconsiders the method of processing images of droplet deposition acquired by low-cost cameras, merging deblurring, super-resolution, and image segmentation to build an end-to-end spray image processor that meets the requirements of rapid in situ spray measurement for scalable application. The specific contributions of this paper are as follows:

1. A droplet image semantic segmentation model is built to output a super-resolution segmentation mask to accurately express the deposited area of individual droplets.
2. The weighted IoU loss function is applied to improve the statistical independence of droplets in super-resolution semantic segmentation.
3. The concave segmentation method is applied for complex droplet adhesion clusters segmentation, including annular type and multiple adhesion.

2. Materials and Methods

### 2.1. Characteristics of Deposited Droplet Images

A portable in situ droplet deposition measurement device was developed to validate the ODEI method. The process of collecting droplet deposition images in the crop canopy using the in situ measurement device is shown in Figure 1. The user controls the acrylic collector to eject to collect spraying droplets with the embedded system, then retracts the collector to capture images. The image processing program embedded in the device analyzes the image and measures the parameters of droplets. In order to measure the deposition of droplets on both sides of the crops' leaves, an LED surface light source is placed above both sides of the collector to obliquely illuminate the front and back sides of the collector where the droplets have been collected. The reflector reflects the images of the front and back sides of the collector to the camera to record a digital image. The optical principles of the device are shown in Figure 2. In practical applications, if the device can be equipped with a camera with lower specifications, the application cost of the method can be reduced, which is conducive to the promotion of scalable application. However, the limitations caused by camera performance in a low-cost imaging system, coupled with a dense spray of droplets and a complex imaging environment, result in blurred and low-resolution images of deposited droplets, which challenges the accurate measurements. The specific characteristics of low-quality images are as follows:

(1) The fog droplet deposition image obtained shows a phenomenon of blurry edges and clear middle.
(2) The blurring at the edge of the image leads to the dissolution of the droplet boundaries in the edge area, and adjacent non adhesive droplets are prone to displaying adhesive properties in the blurred image.

(3) The relatively low resolution cannot accurately show the contour of the edge small droplets, resulting in insufficient accuracy in the statistics for the droplet deposition area.

(4) The ODEI method forms high brightness spots in the center of the deposited droplets, which makes droplet segmentation more difficult.

(5) The surface of the collector shows dirt and scratches after many measurements, resulting in a decrease in the quality of the images collected.
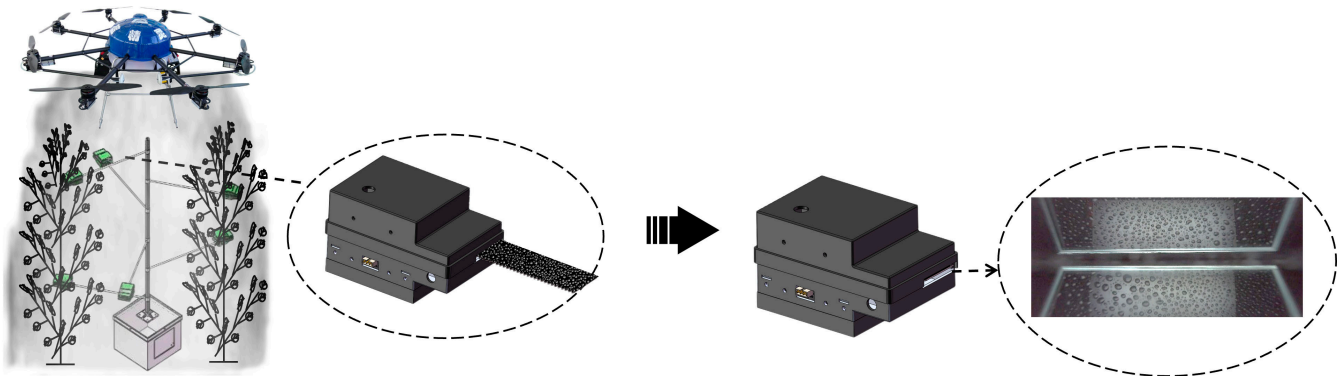
Figure 1. The process of collecting droplet deposition images using in situ measurement devices.
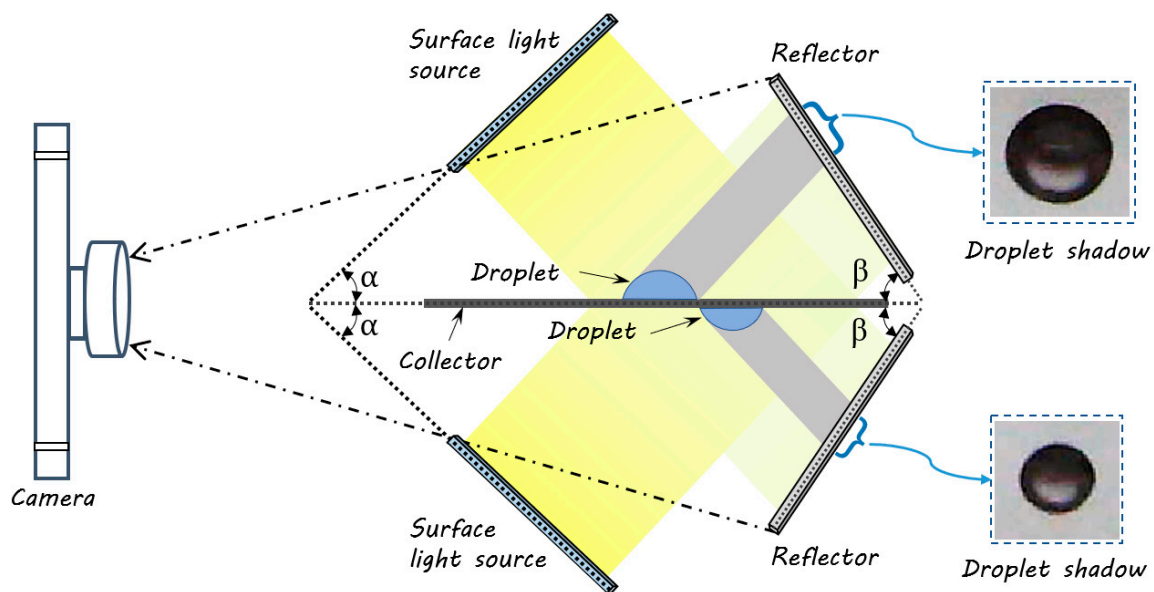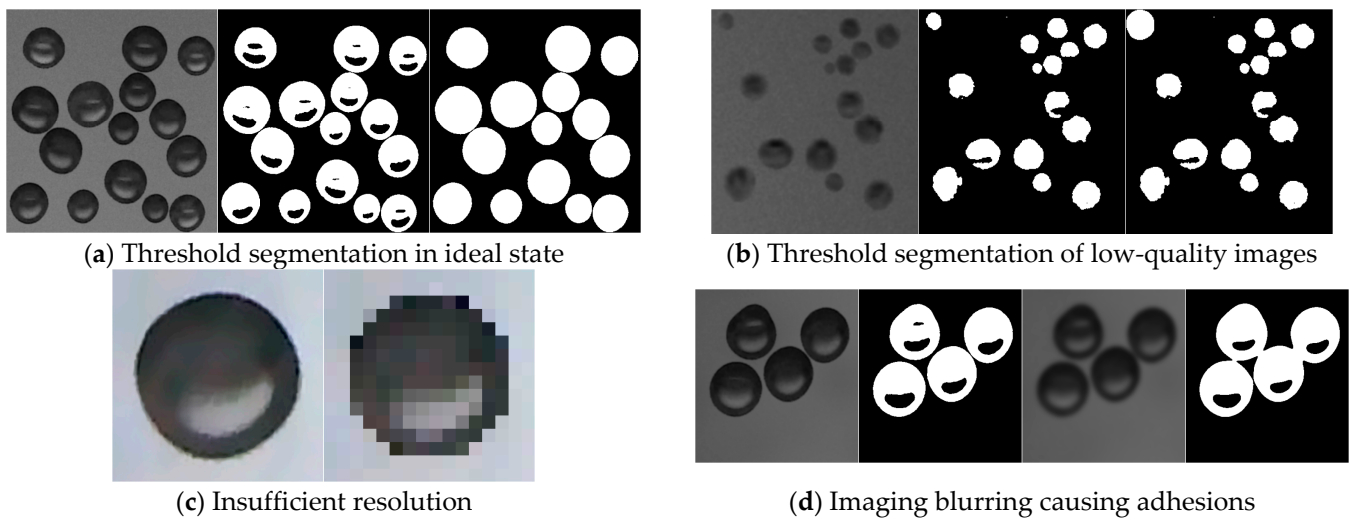

Figure 2. Optical principle of image acquisition device for spray deposition.

The key to spray deposition image processing is to identify individual droplets, calculate their deposition area, convert the deposition area into volume, and then calculate VMD, NMD, and other droplet parameters. As in the device shown in Figure 2, under the illumination of the white LED surface light source, the droplets form an image in the camera as a black ring with a white spot in the center, and the non-droplet area forms an image in the camera with a white background. Calculating the area of individual droplets requires segmenting them from the digital image. Under ideal conditions, the white spot is wrapped by the shadowed edge of the droplet. Although the color and brightness of the white spot and the white background are very close, segmented individual droplets can be achieved by filling them after threshold segmentation, as shown in Figure 3a. However, limited by the performance bottleneck of the camera and some environmental factors, the collected images may be blurry and have insufficient resolution. Severe blur may cause the white spot to deviate from the droplet edge, causing the threshold segmentation and filling method to be ineffective, as shown in Figure 3b. Therefore, accurately obtaining the

area of deposited individual droplets in low-quality deposition images is a challenging task. Insufficient resolution can also lead to inaccurate statistical results for the droplet deposition area. If the number of pixels constituting the droplet is too small, it is hard to accurately characterize the deposition boundary of the droplet, as shown in Figure 3c. Droplet adhesion is another problem that makes droplet area statistics difficult. If the droplets that successively fall on the collector come into physical contact, they will merge into a single droplet and form an elliptical deposition surface under the combined action of surface tension and adhesion. If two droplets are not in physical contact but are very close, image blurring may lead to the edges of the droplets dissolving (blur, spread) and invading each other, forming a connected domain in the image, as shown in Figure 3d, ultimately making it difficult to segment individual droplets. In summary, processing low-quality deposition images is a complex task, and droplet segmentation algorithms need to have the ability to identify individual droplets.


(**a**) Threshold segmentation in ideal state

(**b**) Threshold segmentation of low-quality images

(**c**) Insufficient resolution

(**d**) Imaging blurring causing adhesions

Figure 3. Threshold based droplet segmentation in ideal and actual situations.

2.2. Workflow of Droplet Deposition Low-Cost Measurement

The method proposed in this study aims to construct an image processing process to achieve the measurement of parameters based on low-quality droplet deposition images. Super-resolution semantic segmentation based on deep learning is the core of this method. Datasets, model structures, loss functions, and droplet segmentation algorithms are improved to improve the method's processing capabilities for spray deposition images. The workflow of the method is shown in Figure 4. The blue and green arrows in the figure indicate two processes: the training process and the working process.

Training stage: (1) In order to train the model, image reconstruction-based data augmentation methods were used to create the DDID. The samples included use low-quality images as model input and high-resolution segmentation masks as segmentation annotations. (2) In order to solve the blur problems, insufficient resolution, and imaging adhesion in low-quality droplet deposition image segmentation, a lightweight super-resolution semantic segmentation model DSRSS was constructed to output super-resolution segmentation masks to represent the size and boundaries of individual droplets. (3) In order to improve the droplet independence rate in the DSRSS's output results, a weighted IoU loss function was applied to optimize the category weights, finally achieving high-quality droplet segmentation.
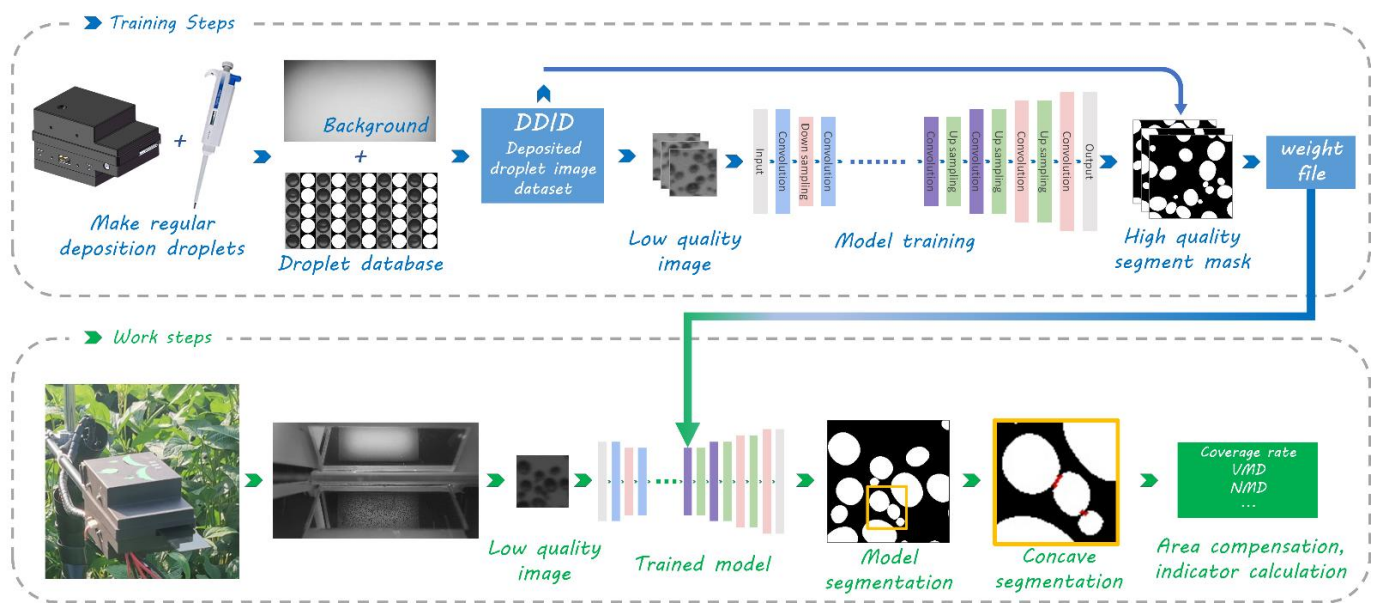
Figure 4. The workflow of parameter acquisition in low-cost droplet deposition measurement.

Working stage: (1) The device collects droplet deposition images and performs pre-processing, including image cropping and distortion correction. (2) The preprocessed image was input into the DSRSS, and through the model's inference computation, a super-resolution segmentation mask end-to-end model was output. (3) The super-resolution segmentation mask is processed with a concave point segmentation algorithm to further improve the independence of droplet segmentation. (4) Droplet measurement indicators, such as coverage rate, droplet density, and VMD, are calculated from the segmentation mask after concave point segmentation.

### 2.3. Construction of Deposition Droplet Image Dataset

2.3.1. Feasibility of Droplet Deposition Image Generation

In order for the model to learn richer morphological patterns of deposited droplets, the annotated GT (ground truth) of the sample must represent the deposition surface of all droplets as accurately as possible, which requires the high-quality image used to annotate the segmentation mask to have a sufficiently high resolution and clarity. An ideal dataset creation method is to use a high-performance camera and a low-performance camera to capture images at the same location. Segmentation masks are annotated with high-quality images taken by high-performance cameras, and they are combined with images taken by low-standard cameras to form training sample pairs. However, it is difficult to capture images with a highly consistent field of view because two cameras with different specifications will generate pose errors, perspective errors, and distortion correction errors during the shooting process, not to mention that the water droplets will evaporate and move during the image switching process. It is difficult to meet these requirements with directly captured images.

According to capillary and wetting phenomena [10], when a droplet with a diameter smaller than the capillary length of the liquid is deposited on a flat solid surface, its water contact angle is constant and appears as a spherical crown with a consistent shape. That is to say, within a reasonable particle size range, taking high-definition images of large droplets and reducing their resolution will result in images that will be highly consistent with directly captured images of small droplets. Therefore, it is feasible to use the data augmentation method of image reconstruction to construct a training data set.

2.3.2. High Quality Image Samples Generation

A large droplet shape dataset containing 67,829 samples was established as the candidate units for generating image samples by extracting individual droplets from high-quality deposition images captured with a high-specification camera. The droplet morphology database includes the grayscale magnification of the droplet region pixels relative to the background pixels in the region and the local segmentation mask of the droplet individual, as shown in Figure 5. The original position of droplets on the collector were also recorded in the database to preserve any correlation with droplet morphology and position. Individual droplets in these databases were used as the foreground during sample generation, with the background coming from empty collector images captured using the ODEI device. During the sample image generation process, individual droplets are scaled to random sizes and attached to the background image according to their original positions recorded in the database, without overlapping each other. Considering that the droplets tilted onto the collector and the droplets which were fused due to physical contact deposit in a long elliptical shape, random shear transformation was also performed on individual droplets. At the same time, the same transformation is applied to the corresponding segmentation mask, and the resulting image sample pair is shown in Figure 6.
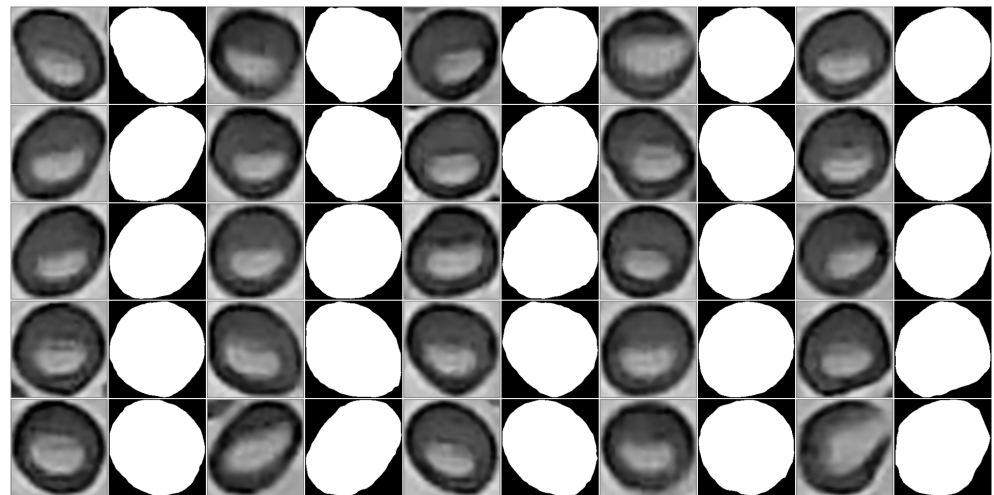
Figure 5. Partial samples from the droplet morphology database.

(**a**) Generated high-resolution image   (**b**) Generated high-resolution segmentation mask
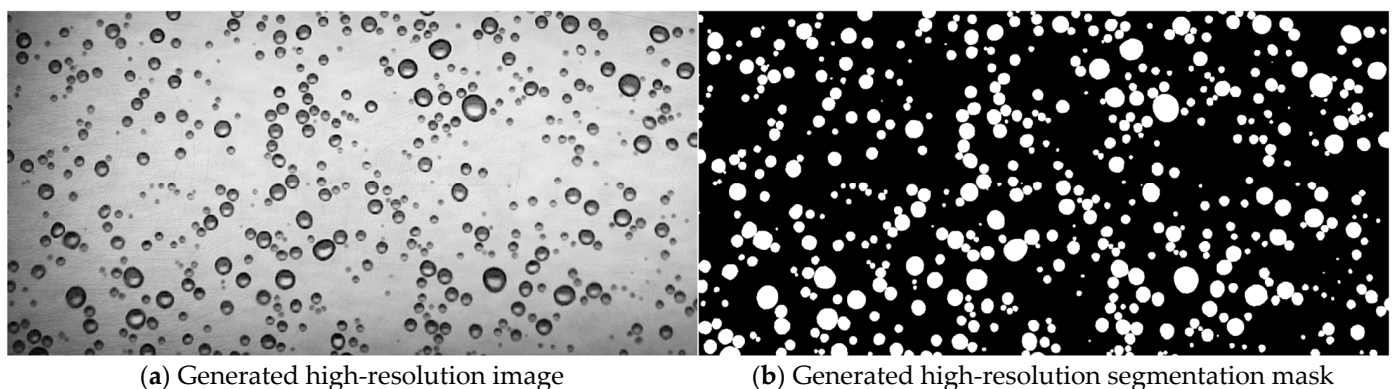
Figure 6. High quality image and synchronously generated segmentation annotations.

2.3.3. Quality-Degrading Operations for the Simulation of Low-Quality Image Acquisition

Low image quality is the result of multiple factors, the most important factors are insufficient COMS sensor performance, which leads to detail loss, and low-performance ultra-wide-angle lenses, which cause spherical aberration. The low-quality images in the dataset are generated from high-quality images through a series of operations, including reducing resolution, adding noise, and applying various fuzzy transformations. Among them,

reducing the resolution is used to simulate the insufficient resolution of low-performance cameras, image noise (including salt and pepper noise and Gaussian noise) and Gaussian blur are used to simulate the loss of image details, and radial blur which gradually increases from the center to the edge of the image is used to simulate spherical aberration, resulting in image margin blurring. Therefore, stains and scratch textures were randomly added to the low-quality images to simulate real images collected by the device in field applications.

Since scratches are physically present textures on the collector, the scratch simulation in the generated image is directly performed on the high-resolution image background, prior to the attachment of droplets. The scratch simulation has actually been completed in the generated image shown in Figure 6a. Spherical aberration is caused by the abnormal refraction of light by the lens, which occurs before the imaging of CMOS sensors, so the application of radial blur is immediately followed by scratch simulation. Gaussian blur, adding noise, and reducing resolution are all implemented to simulate image detail loss. However, Gaussian blur reduces the effect of salt and pepper noise, and reducing resolution weakens the effect of Gaussian blur. Therefore, the order of executing these three steps is: reducing resolution, Gaussian blur, and adding noise. The simulation process for the low-quality images is shown in Figure 7, and a comparison between the generated low-resolution images and real low-quality images is shown in Figure 8.
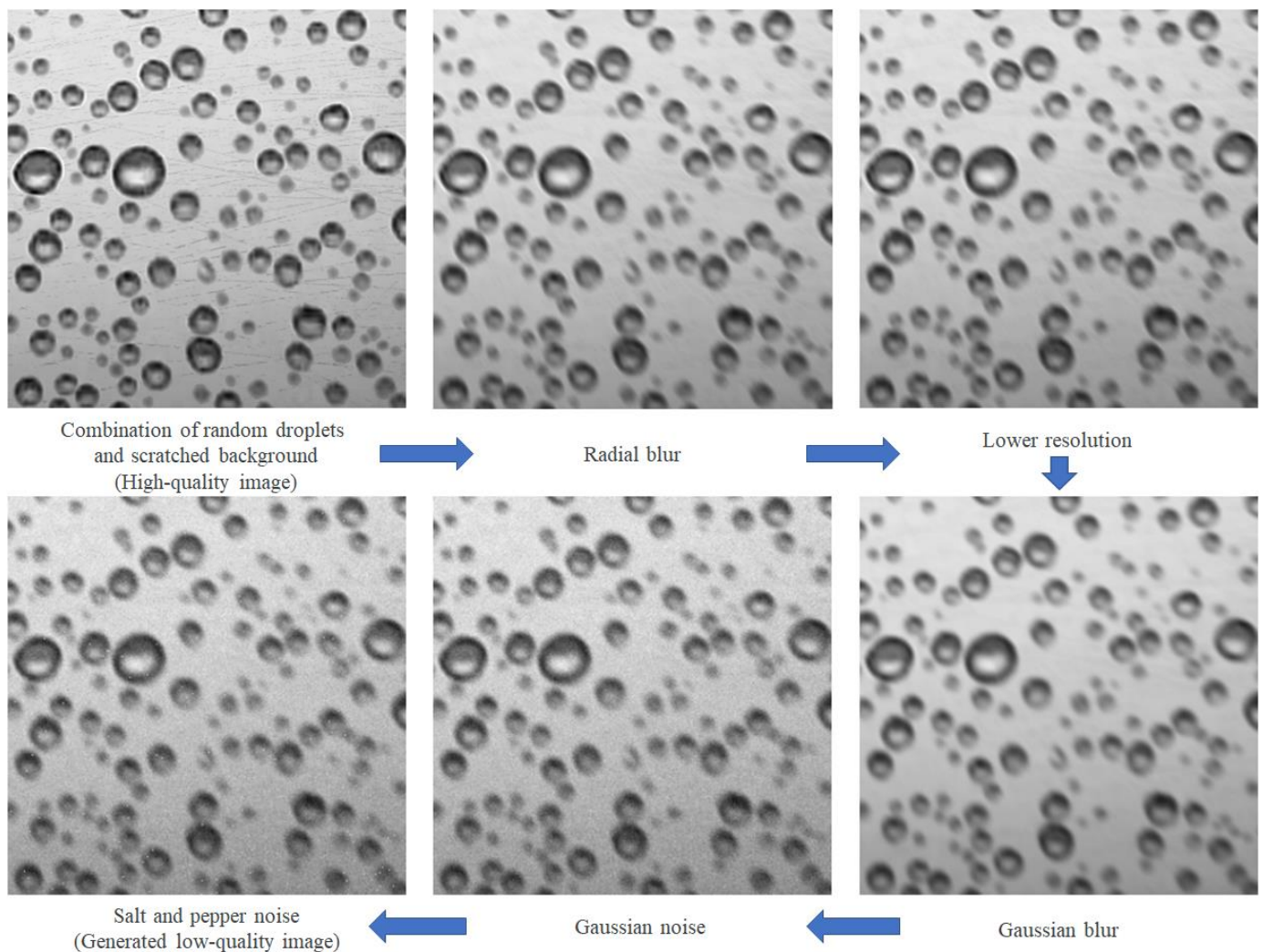
Figure 7. The simulation process for low-quality images.

(**a**) Real low-quality image   (**b**) Generated simulated low-quality image
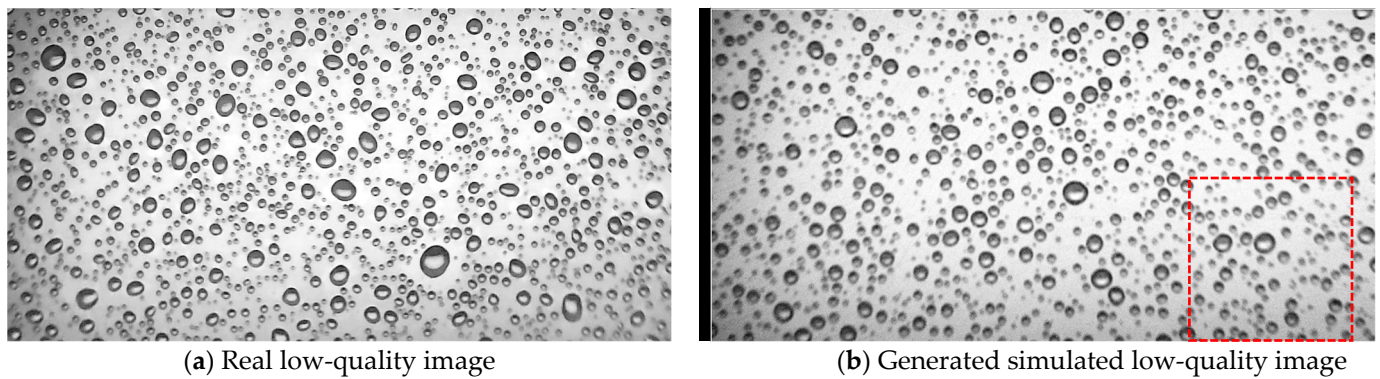
Figure 8. Low quality image samples. (The part in the dotted box is the area shown in Figure 7).

In order to meet the needs of model training, validation, and comparison, each group of samples in the dataset also contains additional sub items, which come from the intermediate process of dataset generation. Finally, a total of 2000 groups of training samples were generated, of which 1600 groups were used as the training set and 400 groups were used as the validation set. In this paper, the above dataset was named DDID (deposited droplet image dataset).

2.4. Droplet Super-Resolution Semantic Segmentation (DSRSS) Method

2.4.1. Structure of the DSRSS

In the field of deep learning image segmentation, researchers use large-scale natural image datasets to verify the versatility and generalization ability of the model. The scale of SOTA models is getting larger and larger, including an increasing number of special structures, which also puts forward increasing requirements for computer performance. Compared with natural image datasets, the information contained in spray deposition images is relatively simple. Applying large models to solve the segmentation of droplet images will cause computer hardware resources to be wasted and is not conducive to cost control. Due to the above reasons, this study built a lightweight CNN model to meet the requirements for application in embedded systems.

The idea of super-resolution was used to solve the problem of insufficient resolution of small droplets. DSRSS is a super-resolution semantic segmentation model used for dealing with low-quality droplet deposition images. The length and width of the output segmentation mask are four times that of the input image, which can more accurately express the morphology of the outline of the droplets. In common super-resolution image reconstruction tasks [14,16], the deep learning model learns the prior knowledge contained in large-scale datasets during training and applies it to the restoration of image details such as color and texture, which places higher requirements on the number of parameters in the model. The super-resolution semantic segmentation task does not consider complex image details and only focuses on outputting more accurate segmentation masks. It has lower requirements for the number of model parameters and is conducive to lightweighting the model.

DSRSS adopts an encoder–decoder structure, which is widely used in semantic segmentation models, such as U-Net [17] and its derivative models, SegNet [18], DeepLab V3+ [19], etc. The function of the encoder is feature extraction, and the decoder generates a segmentation mask based on the features. Under the guidance of the training data, the decoder can learn some prior knowledge and act on image reconstruction.

DSRSS is a fully convolutional network that does not contain fully connected layers [20]. It uses standard 2D convolution operations as convolutional layers, and batch normalization and swish [21] function activation are performed after each convolutional layer. On the encoder side, the activation function is followed by a pooling operation with a kernel size of 2×2, and on the decoder side; a bilinear interpolation up-sampling operation with a factor of 2×2 is performed before each convolutional layer. The cross-layer connection

structure in U-Net was referenced. The feature map obtained by each down-sampling in the encoder is concatenated with the feature map of the same level in the decoder. In the test, this structure can significantly improve the *IoU* of the model output.

The structure of the DSRSS model has been determined through a large number of experiments, with its depth and width at relatively optimal values. Cross layer connection structure is also essential. The down-sampling level must reach at least 3 to ensure the high performance of the model. If the down-sampling level is lower than 3, errors in the segmentation of droplet spots are prone to occur. To achieve 4× super-resolution output, the number of up-sampling operations should be 2 more than the number of pooling operations. The width of the model also has a great impact on the performance of the model, but the larger the scale of the model, the higher the requirements for hardware performance. The final model structure adopted is shown in Figure 9. On this basis, the benefits of continuing to increase the width of the model are no longer clear.
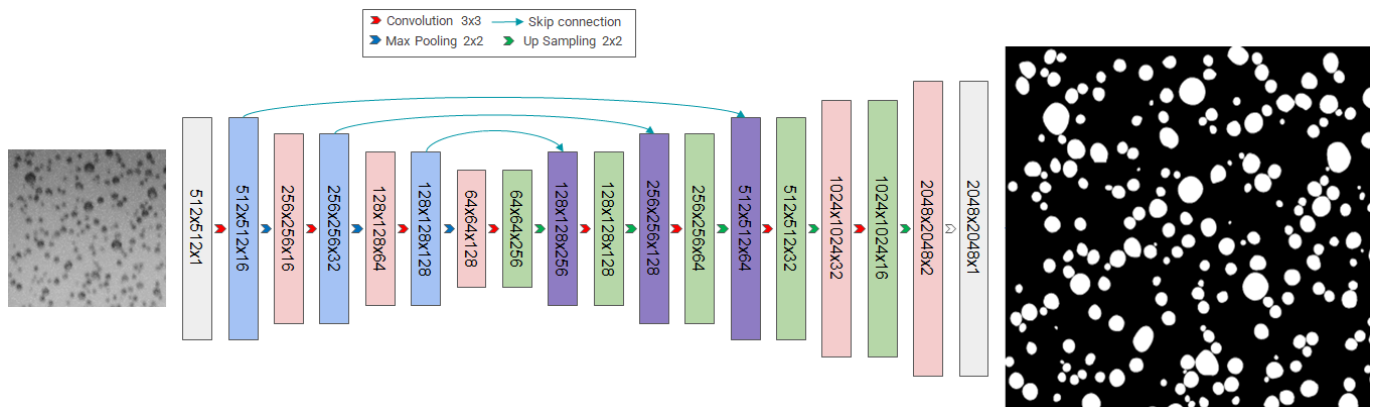
Figure 9. Structure of DSRSS.

2.4.2. Loss Function and Metrics

The independence of individual droplets and the accuracy of segmentation areas are the two most important criteria in measuring the quality of DSRSS segmentation, and the loss function determines the model's optimization goal. Semantic segmentation does not pay attention to the independence of individual droplets, and its optimization is usually directed towards improve the accuracy of segmented regions. In order to enable the model to take into account the independence of individual droplets, specific metrics and loss functions were applied.

Three normalization metrics, including the recognition rate, independence rate, and correction rate, of individual droplets were used to characterize the independence of individual droplets. The calculation method is: (1) Search for contours in the segmentation mask output by the model to obtain a list of predicted droplets. (2) Search for contours in GT to obtain a list of target droplets. (3) Traverse the target droplet list, count the proportion of individuals that only intersect with one predicted droplet and this intersecting predicted droplet has no other intersecting droplets as the independence rate, and count the proportion of individuals that intersect with any predicted droplet as the recognition rate. (4) Traverse all predicted droplets and count the proportion of individuals that intersect with any target droplet as the accuracy rate. The three quantity standards can be expressed in Equation (1).

$$\begin{cases} T_r = \dfrac{N_{true}}{N_t} \\[2mm] I_r = \dfrac{N_{indep}}{N_t} \\[2mm] C_r = \dfrac{N_{true}}{N_p} \end{cases} \tag{1}$$

where, T_r denotes the recognition rate, I_r denotes the independence rate, C_r denotes the correct rate, N_t denotes the number of target droplets, N_p denotes the number of predicted droplets, N_{true} denotes the number of correctly recognized droplets, and N_{indep} denotes the number of droplets without adhesion.

The *IoU* is often used as one of the evaluation criteria for semantic segmentation and is also used to define the loss function. The *IoU* loss function optimizes the intersection area and union area ratio of the target area and the prediction area toward 1, so that the segmentation map output by the model is as consistent as possible with GT. The *IoU* loss function and its difference set form a two-category segmentation which can be expressed by Equation (2).

$$L_{IoU} = 1 - \frac{A_t \cap A_p}{A_t \cup A_p} = \frac{(A_t \backslash A_p) + (A_p \backslash A_t)}{A_t \cup A_p} \tag{2}$$

where A_t denotes the target area, A_p denotes the prediction area, and L_{IoU} denotes the loss value.

In the segmentation mask output by the model, the droplet adhesion points are composed of very few pixels. However, the segmentation of these pixels determines the model's ability to segment individual droplets, so their segmentation should be of higher priority. This special priority can be reflected in the loss function. Since the adhesion points between droplets are only included in the difference set $A_t \backslash A_p$ between the prediction area and the target area, reducing the proportion of this difference set in the loss function can adjust the sensitivity of the model to droplet adhesions. So, two weight values α and β were added into Equation (2), and the loss value is represented by the symbol L_{IoU}. The modified loss function was named *WIoU*, as shown in Equation (3).

$$L_{WIoU} = \frac{\alpha * (A_t \backslash A_p) + \beta * (A_p \backslash A_t)}{A_t \cup A_p} \tag{3}$$

The imbalance of weights α and β will change the optimization direction of the model, improve the independence rate, and lead to a decrease in *IoU* metric. Precision and Recall can reflect the above influence. These two evaluation indicators can be represented with Equation (4), and the meaning of the symbols is the same as Equation (1).

$$\begin{cases} Precision = \dfrac{A_t \cap A_p}{A_p} \\ Recall = \dfrac{A_t \cap A_p}{A_t} \end{cases} \tag{4}$$

The *IoU* can be regarded as a combination of *Precision* and *Recall*. Theoretically, the larger the weight α and the smaller the weight β in Equation (3), the higher the Precision of the model and the lower the probability of adhesion; however, the Recall will decrease. Therefore, there must be an optimal combination of α and β that can balance the independence of the droplets and the accuracy of the segmentation area to achieve the optimal segmentation effect.

The sum of a total of six indicators including *IoU*, *Precision*, and *Recall*, which characterize the accuracy of the region, and T_r, I_r, and C_r, which characterize the independence of the droplets, represented by the symbol M, was used as the criterion for judging the quality of the model segmentation effect to estimate the optimal weight coefficients α and β, as shown in Equation (5).

$$M = IoU + Precision + Recall + T_r + I_r + C_r \tag{5}$$

2.5. Multi-Adhesion Concave Segmentation (MACS) Algorithm

The output of the DSRSS model does not contain individual information, and additional algorithms are required to extract individual droplets and calculate their parameters.

Although the droplet independence rate in the DSRSS model's output is already very high after WIoU loss function optimization, it cannot reach 100%, especially when processing extremely low-quality images. Therefore, when extracting individual droplets, it is necessary to further segment the adherent droplets.

Contours were used to represent individual droplets, and the roundness of the contour is used to determine whether there is adhesion in the extracted droplet patches. Contours with a roundness greater than or equal to the threshold are judged as independent droplets, otherwise they are judged as adherent droplets. The formula for calculating roundness is shown in Equation (6).

$$R = \frac{4\pi A}{L^2}, 0 < R < 1 \tag{6}$$

where A denotes the area of the segmented patch, L denotes the perimeter of the segmented patch, and R denotes the roundness of the segmented patch.

The algorithm for separating adherent droplets is the concave point method, which has many implementation methods. The most critical step in the concave point method is to locate the concave points in the image. In the high-resolution segmentation mask output by the DSRSS model, the adhesion point of the droplet has a sharp contact angle, so the concave point of the adhesion point can be regarded as a corner point and located through the contour corner detection algorithm. Another key step in the concave segmentation method is the pairwise matching of concave points. Since the deposition bottom surfaces of droplets are all approximately elliptical, the slope of a straight line passing through a pair of concave points must be between the slope of the straight line connecting the deepest point of the concave point and its two adjacent points before and after it. The distance between two concave points of the same adhesion is shorter than the distance from them to other concave points that satisfy the slope condition. Based on the above principles, the pseudo code of Concave Segmentation is shown in Algorithm 1.

Figure 10 shows the segmentation process for a complex adhesion cluster. After segmentation, the contour list of individual droplets can be directly obtained using the *findContours* and *contourArea* functions in *OpenCV*. Then, the deposition area of individual droplets in the contour list can be calculated and converted into volume. Further calculations can produce complex droplet measurement indicators such as *VMD* and *NMD*.

(a)   **(b)**   **(c)**   **(d)**
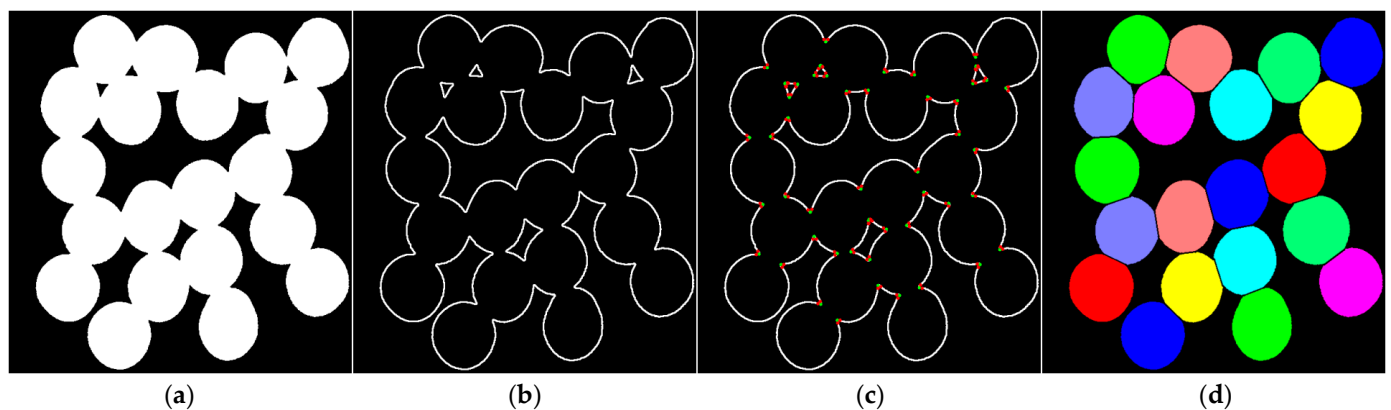
Figure 10. Applicable scenarios in and effects of concave segmentation. (**a**) Complex adhesion patch. (**b**) Find contours. (**c**) Locate the concave point. (**d**) Concave point matching and segmentation (Colors are used to distinguish individuals).

Algorithm 1 Concave Segmentation algorithm

Input: Droplet adhesion patch image $I_{adherent}$
Output: Droplet individual patch image $I_{individual}$
Search for contours in droplet patch images
$L_{contours} \leftarrow$ findContours($I_{adherent}$)
Obtain corner point list using contour corner detection
$L_{corners} \leftarrow$ detectCorners ($L_{contours}$)
Locate the corner point and its two adjacent points before and after in the outline, and obtain
3 coordinate arrays.
L_{corner_locs}, L_{prev_locs}, $L_{next\_locs} \leftarrow$ getLocations ($L_{corners}$)
Concave point pairing.
completed_pnts $\leftarrow$ [] # Record the paired droplets
$I_{individual} \leftarrow I_{adherent}$
for corner_loc_1, prev_loc_1, next_loc_1 $\in$ (L_{corner_locs}, L_{prev_locs}, L_{next_locs}) do
       if corner_loc_1 $\in$ completed_pnts then
          continue
       min_dist_corner_loc = null
       min_distance $\leftarrow \infty$
       for corner_loc_2, prev_loc_2, next_loc_2 $\in$ (L_{corner_locs}, L_{prev_locs}, L_{next_locs}) do
          if corner_loc_2 $\in$ completed_pnts then
             continue
          dintance $\leftarrow$ getDistance(corner_loc_1, corner_loc_2)
          if dintance **<** min_distance then
             if line_through_triangle(corner_loc_1, prev_loc_1, next_loc_1,
                           corner_loc_2, prev_loc_2, next_loc_2) then
                min_dist_corner_loc $\leftarrow$ corner_loc_2
                min_distance $\leftarrow$ dintance
       if min_dist_corner_loc is not null then
          completed_pnts $\leftarrow$ completed_pnts $\cup$ corner_loc_1
          completed_pnts $\leftarrow$ completed_pnts $\cup$ min_dist_corner_loc
          line($I_{individual}$, corner_loc_1, corner_loc_2)
    return $I_{individual}$

2.6. Droplet Segmentation Area Correction and Deposition Parameter Acquisition

Calculating *VMD* and *NMD* requires traversing individual droplets, so identifying individual droplets is the most basic prerequisite for calculating these two indicators. When the independence rate is regarded as the most important metric in the model, the accuracy of the droplet segmentation area may not be in the optimal state, resulting in statistical deviations in *VMD*, *NMD*, and global coverage. Therefore, it is necessary to correct the inaccurate segmentation area results.

The semantic segmentation model is equivalent to a function that implements mapping the image to the segmentation mask. It contains the rules for segmenting droplets with the model, which will be reflected in the segmentation results of the model. Finding this rule can correct the segmentation area of individual droplets, and improve the effectiveness of spray evaluation indicators such as *VMD*, *NMD*, and global coverage to a certain extent.

In the practical application of the method, the predicted area of the droplets is the only known quantity. If there is a mapping relationship between the predicted area A_{pred} of the droplet and the real area A_{true} of the droplet, deviation correction can be directly implemented through this relationship. According to experimental statistics, this relationship is not a fixed ratio, but changes with the size of the droplet. The smaller the size of the droplet, the greater the ratio $\gamma = \dfrac{A_{pred}}{A_{true}}$. When the droplet reaches a certain level, γ approaches a certain value. In addition, if the relationship between A_{pred} and A_{true} is directly fitted, the deviation in the area of small droplets will be ignored. The range of the area ratio γ's value is [0, 1], which is more accurate for small droplets when fitting the relationship and is more suitable as the dependent variable in the relationship function. If $\gamma = f_{\gamma}(A_{pred})$ is used to

represent the mapping function from the predicted area of the droplet to the area ratio γ, the deviation value can be expressed by Equation (7).

$$\delta = \frac{A_{pred}}{f_\gamma\left(A_{pred}\right)} - A_{pred} \tag{7}$$

The individual area of droplets corrected by Formula (6) is used to calculate droplet deposition parameters, including droplet density (D_t), droplet coverage (Dr), *VMD*, *NMD*, etc., which can be represented by Equations (8)–(11).

$$D_t = \frac{n}{S_{img}} \tag{8}$$

where n denotes the number of droplets in the sample, and S_{img} denotes the total area of the spray deposition image.

$$D_r = \frac{\sum_{i=1}^{n}\left(\frac{A_i}{f_\gamma(A_i)}\right)}{S_{img}} \tag{9}$$

where n denotes the number of droplets in the sample, A_i denotes the predicted area of the i-th droplet, f_γ denotes the correction formula from the predicted area to the real area, and S_{img} denotes the total area of the spray deposition image.

$$VMD = \sqrt[3]{\frac{6 * f_{vd}(L_{ads})}{\pi}} \tag{10}$$

where, L_{ads} denotes an array of area values sorted from small to large for all droplets in a single image sample, while f_{vd} denotes a function that includes logical judgment. It converts the input ordered area array into a volume sequence and calculates the volume of the droplet when the cumulative values of the elements in the volume sequence from small to large reach 50% of the total volume sequence.

$$NMD = \sqrt[3]{\frac{6 * f_{nd}(L_{ads})}{\pi}} \tag{11}$$

where L_{ads} denotes the area values of all droplets in a single image sample sorted from small to large, while f_{nd} denotes the volume calculated from the median element of the area array.

3. Results and Discussion

3.1. Results

3.1.1. Semantic Segmentation Result of DSRSS Model

In order to determine the optimal α and β in the *WIoU* loss function, the model was trained on the DDID under different α and β values. The convergence curve of the model is shown in Figure 11. The performance indicators that the model can achieve when training for 100 epochs are shown in Figure 12. Each subgraph represents the optimal values that the model can achieve for each metric when training for 100 epochs under a pair of parameters α and β. The more balanced these optimal values are, the larger the area of the corresponding hexagonal shape enclosed by the points, indicating the better performance of the model. Figure 12 shows that the overall performance of the model reaches its optimum when α and β are 0.775 and 0.225, respectively. At this time, the *IoU*, recognition rate T_r, independence rate I_r, and correct rate C_r of the model are relatively high, and Precision and Recall are relatively balanced.
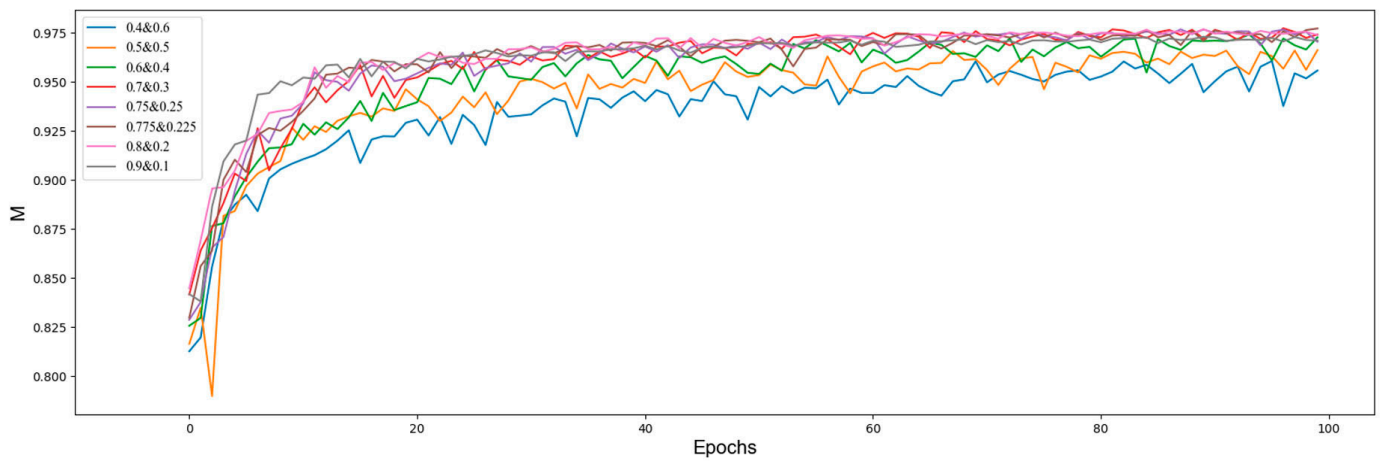
Figure 11. Convergence curve of the model under different weight values of the *WIoU* loss function (*α* and *β*, 100 epochs).
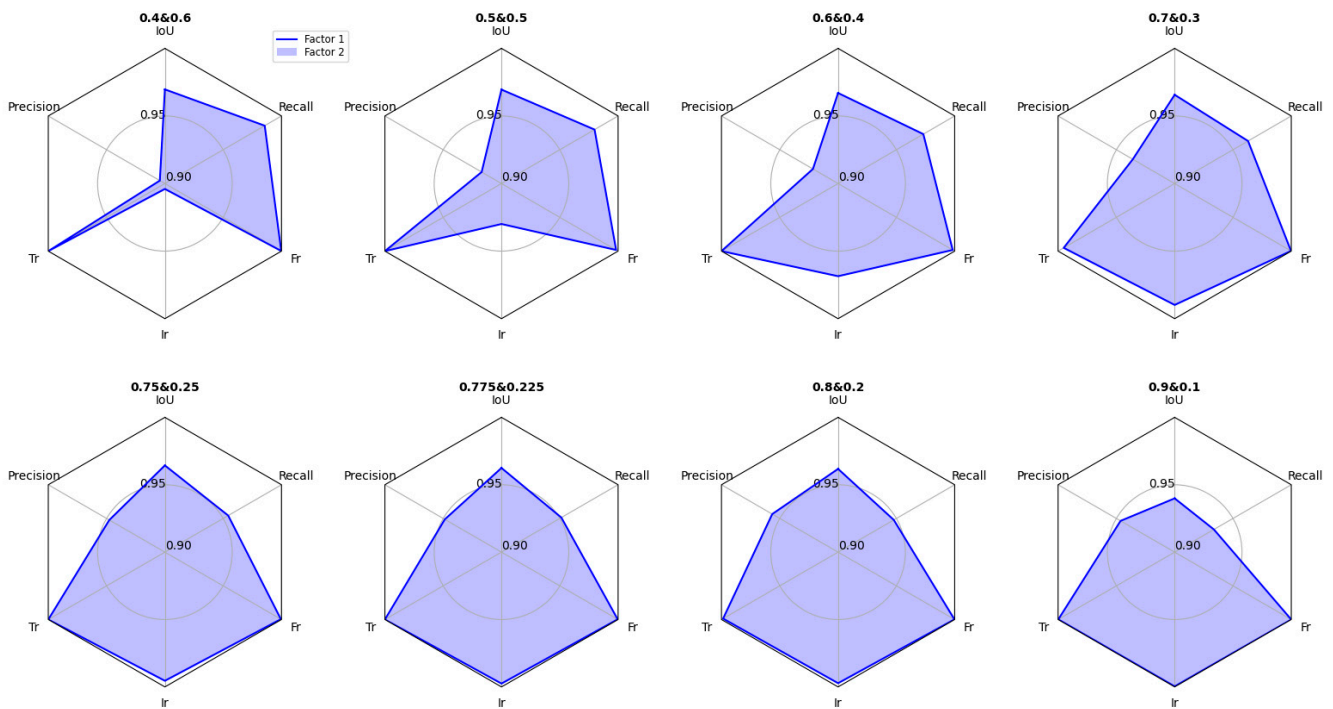
Figure 12. Optimal performance of the model under different weight values of the *WIoU* loss function (*α* and *β*, 100 epochs).

After determining the optimal α and β, the DSRSS model was trained more fully. The *IoU*, *Precision*, *Recall*, *Tr*, *Ir*, and *Cr* change curves of the model are shown in Figure 13. It shows that the convergence of *IoU*, Precision, and Recall is very smooth, while the convergence of *Tr*, *Ir*, and Cr is more turbulent, and the independence rate oscillates far more than other indicators. Therefore, the independence rate is used as the only indicator of the optimal weight value. In 1000 epochs of training, the independence rate of the model on the verification set reached 0.998, the recognition rate and correct rate both reached close to 1.0, and the *IoU* reached 0.973. Figure 14 shows the model's segmentation effect on the DDID verification sample. The comparison of the sample image with the colored segmentation mask shows that the segmentation of the edge of droplet is accurate, but the droplet independence needs to be improved.
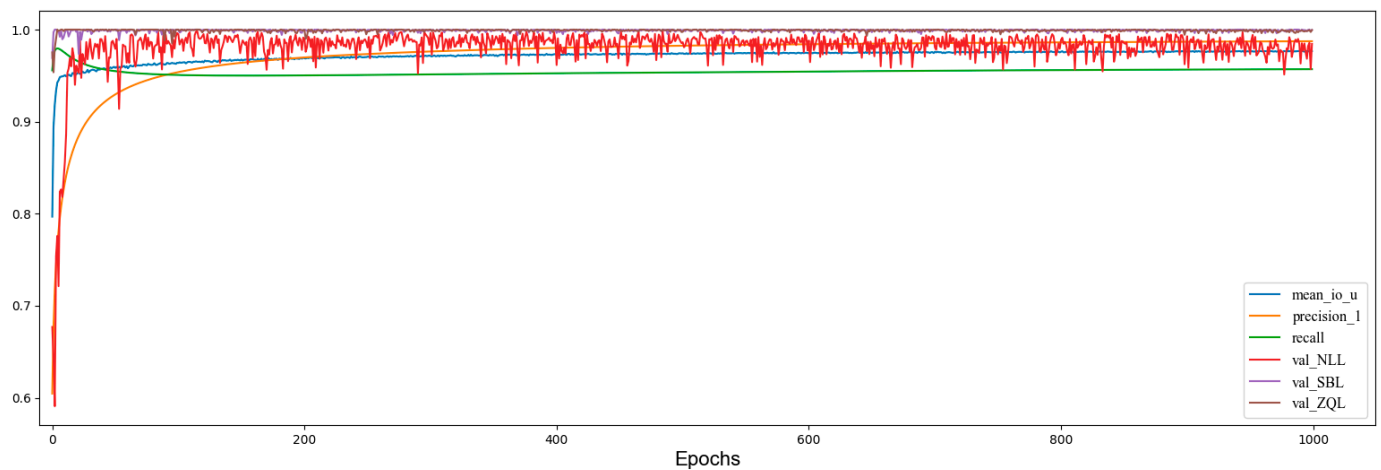
Figure 13. Full convergence of the model.

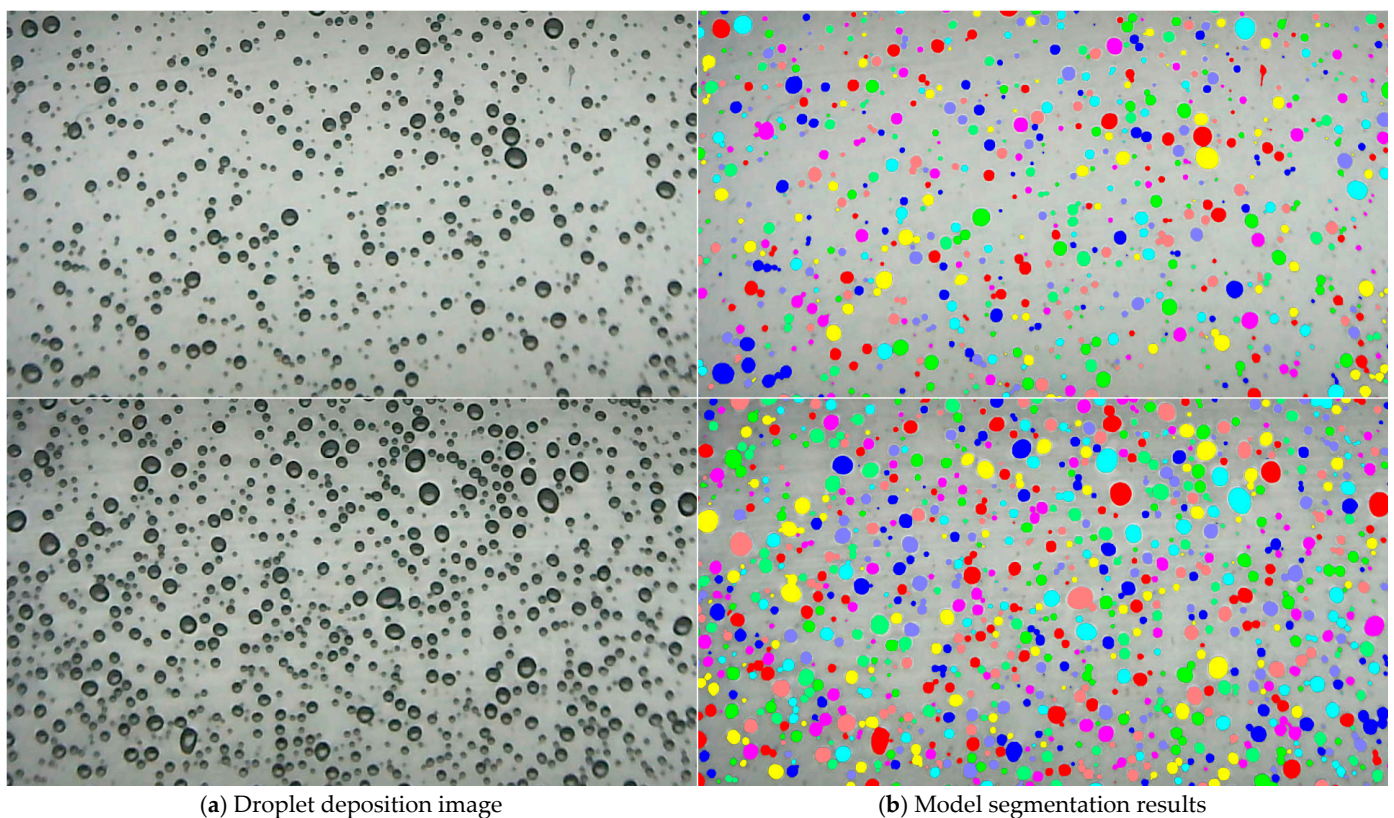(**a**) Droplet deposition image　　　　　(**b**) Model segmentation results

Figure 14. Segmentation results for fully trained model.

3.1.2. Adhesion Segmentation Result of MACS Algorithm

In order to verify the effect of the MACS algorithm, samples with poor quality, dense droplets, and severe adhesion were selected from the validation dataset set for testing. An example of the selected image sample is shown in Figure 15a. The results show that the auxiliary segmentation algorithm can further improve the independence rate of droplet segmentation. The average independence rate of 22 samples increased by 15.7%, as shown in Figure 15b.
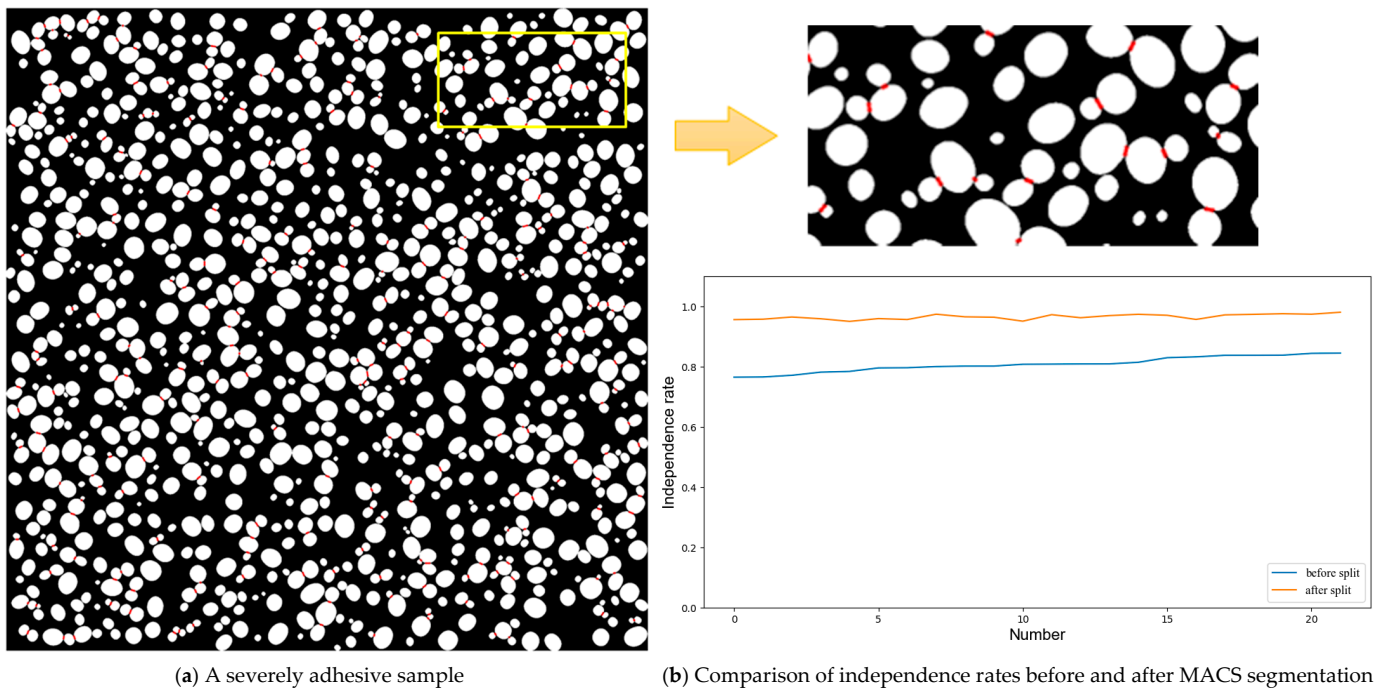
(**a**) A severely adhesive sample   (**b**) Comparison of independence rates before and after MACS segmentation

Figure 15. Testing of MACS algorithm.

3.1.3. Result of Droplet Segmentation Area Correction

The droplet areas predicted by the model for the validation dataset and the ratio of these predicted areas to the corresponding labeled areas were counted. After removing incorrectly segmented droplets and adherent droplets, a total of 98,795 pairs of original data were obtained. The distribution of the predicted droplets was divided into 256 intervals, and all data in each interval were averaged, resulting in a total of 256 pairs of mean sample data. Using 256 pairs of mean data, the conversion formula Equation (12) was fitted. The R2 value of Equation (12) and 256 mean sample points reached 99.07. The distribution of original sample pairs, the distribution of mean sample pairs, and the graph of Equation (12) are shown in Figure 16.

$$\gamma = \frac{1}{\left(1 + 24.0717268^{-A_{pred}^{0.133783199}}\right)^{73.9259088}} - 0.0196201541 \tag{12}$$
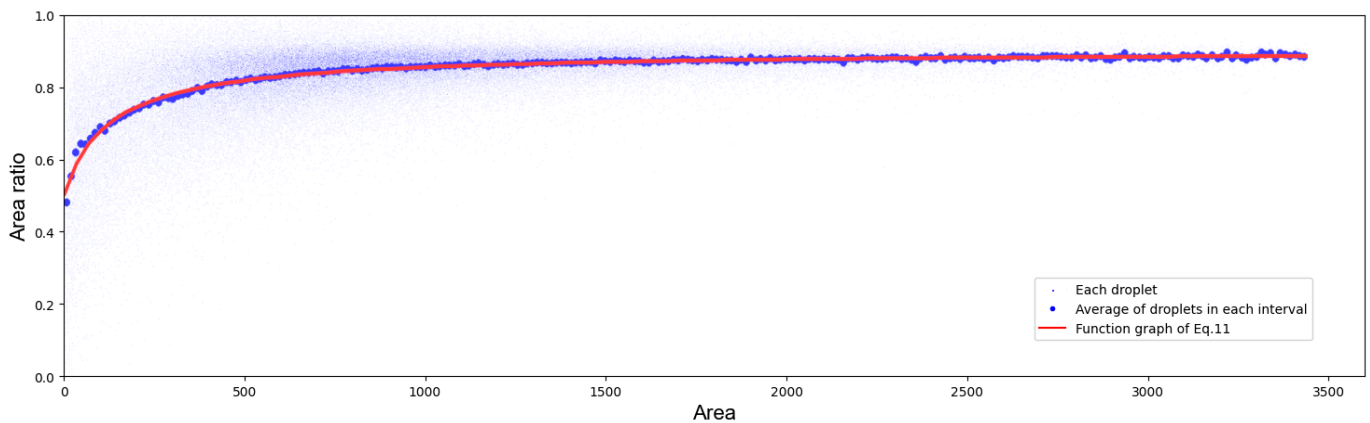

Figure 16. γ Distribution and Function Fitting.

In order to verify the effect of Equation (12) in global indicator statistics, the validated samples were selected to calculate three coverage rates. They are the ground truth coverage,

the coverage directly calculated from the model output, and the model's output coverage corrected by Equation (12). Each group of data were arranged from large to small according to the ground truth coverage, and the line chart drawn is shown in Figure 17. The correction function reduces the coverage error from 4.67% to 1.12%, showing that the correction function can improve the global statistical indicators.
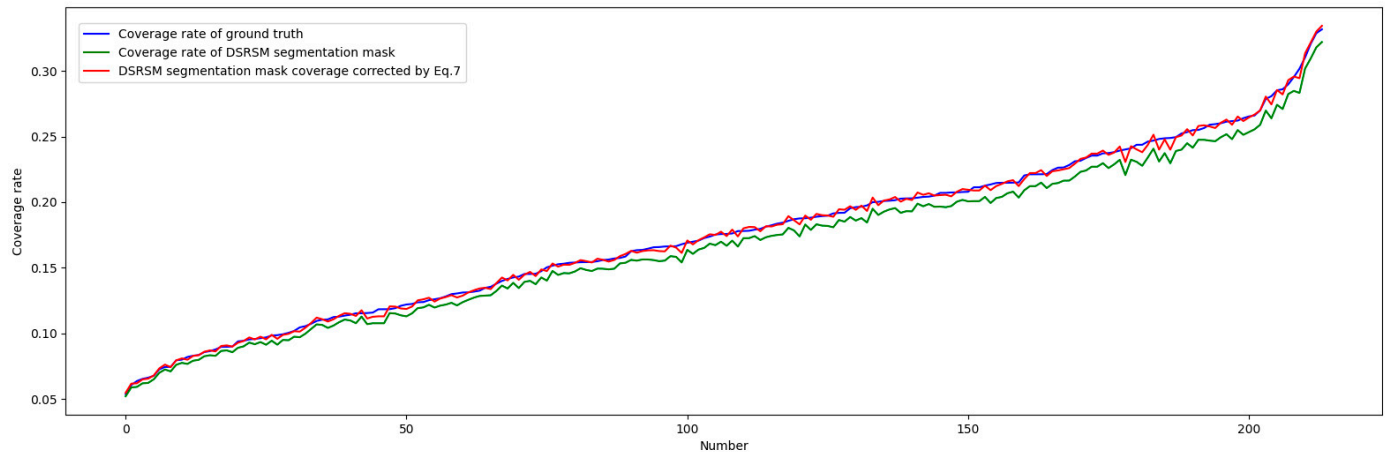
**Figure 17.** The effect of correction function on coverage rate.

3.1.4. Result of Real Sample Processing

Finally, we input the real droplet deposition images captured by the device into the model, and the segmentation mask obtained after inference is shown in Figure 18. Perceptually, the segmentation performance of the model is significantly better than that of the adaptive threshold segmentation (ATS) method. The vast majority of droplets are effectively segmented, including those that are adhered during threshold segmentation, and the obtained edges are smoother than threshold segmentation, filling almost all droplet bright spots. The above advantages are also reflected in the spray parameters; Table 1 shows the spray parameters calculated based on the segmentation masks output by the two methods, including the coverage, droplet number, droplet density, and VMD. Comparing the segmentation boundaries of droplets in the two segmentation masks in Figure 18, the coverage obtained by our method is closer to the real situation, and due to the higher droplet independency rate, our method outputs a higher droplet density and a smaller value VMD value.
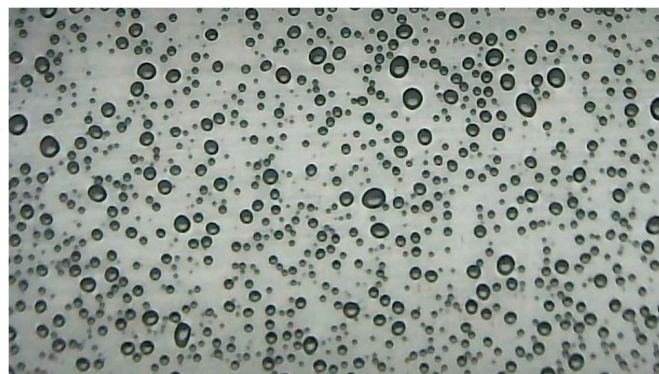
(**a**) Real droplet deposition image

**Figure 18.** *Cont.*

(**b**) ATS method

(**c**) Our method
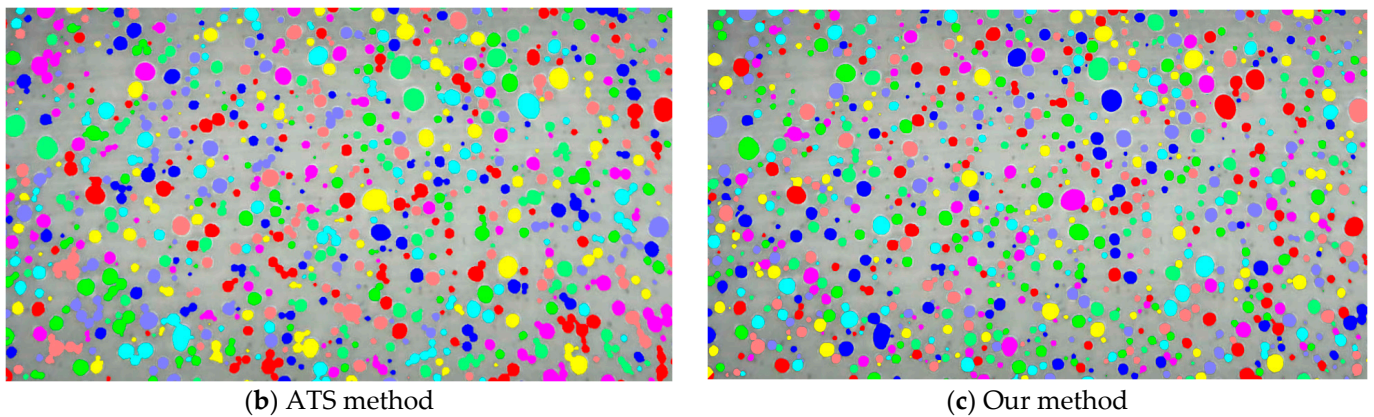
Figure 18. Comparison of performance between DSRSS and ATS methods and the real droplet deposition image.

Table 1. Comparison of spray parameters between DSRSS and ATS methods.

Segmentation Method	Coverage Rate	Number of Droplets	Droplet Density (n/mm^2)	VMD (µm)
ATS	0.220	835	202	98.1
Our method	0.176	1044	252	76.8

3.1.5. Ablation Experiment of the Structure of the DSRSS Model

To verify the necessity of each component in the model, ablation experiments were performed. Figure 19 shows the performance of the model under four conditions: unchanged, downsampling reduced by one level, halved width, and cross layer connection removed. CLearly, the default structure model has the better performance.
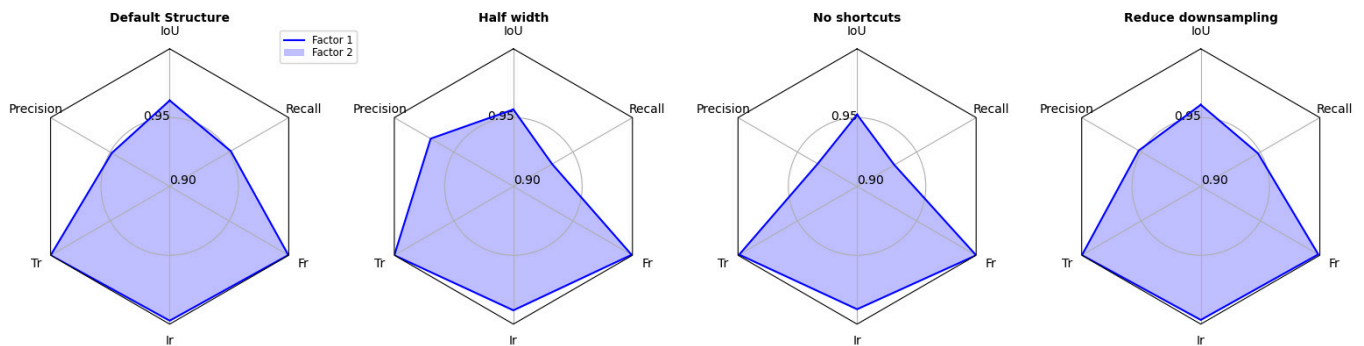
Figure 19. Performance of the model with changed structure.

3.2. Discussion

3.2.1. Rationality of Semantic Segmentation

From the perspective of the form of the task, target detection, instance segmentation, and semantic segmentation based on deep learning can all be applied to the segmentation and statistics of droplet images, replacing some processes in spray measurement work, as shown in Figure 20. The colored parts in the figure represent the work that the corresponding method can accomplish.


Figure 20. The role of three types of depth learning methods in spray measurement.

Among the above three types of methods, the instance segmentation algorithm can directly output the mask of individual droplets and realize the direct calculation of the deposition area, which seems to be the most reasonable choice. However, instance segmentation of both small objects and dense objects is extremely challenging. Perhaps because the droplets of high-quality spray are too small and too dense, in our tests, several instance segmentation models failed to show an acceptable segmentation accuracy, as shown in Figure 21.

(**a**) (**b**) (**c**) (**d**)
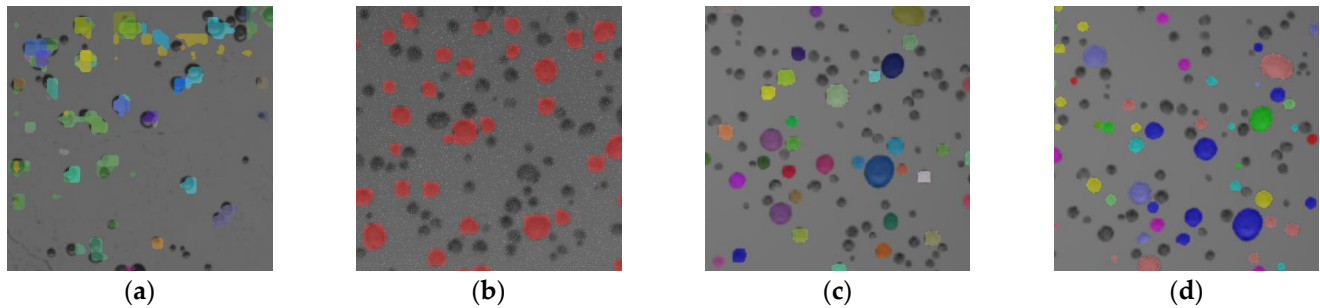
Figure 21. Performance of four instance segmentation methods in droplet segmentation. (**a**) YOLACT [22]. (**b**) YOLOv8 [23]. (**c**) FastSAM [24]. (**d**) SAM [25].

Since the segmentation task only needs to segment a single type of object, all other areas in the image except the droplets are considered as background. First, we performed semantic segmentation on the image, obtained the segmentation mask, and then segmented the individual droplets through an additional process, which is equivalent to using the instance segmentation task on a single type of object. Moreover, DSRSS already has a certain ability to focus on individual droplets. If the image is enlarged so that the size of the droplets is far beyond the reasonable range, and then input into the model, the result shown in Figure 22 is obtained. This shows that the display model is able to "consciously" reconstruct the shape and contour of the droplet from the blurred image based on the prior knowledge it learned.
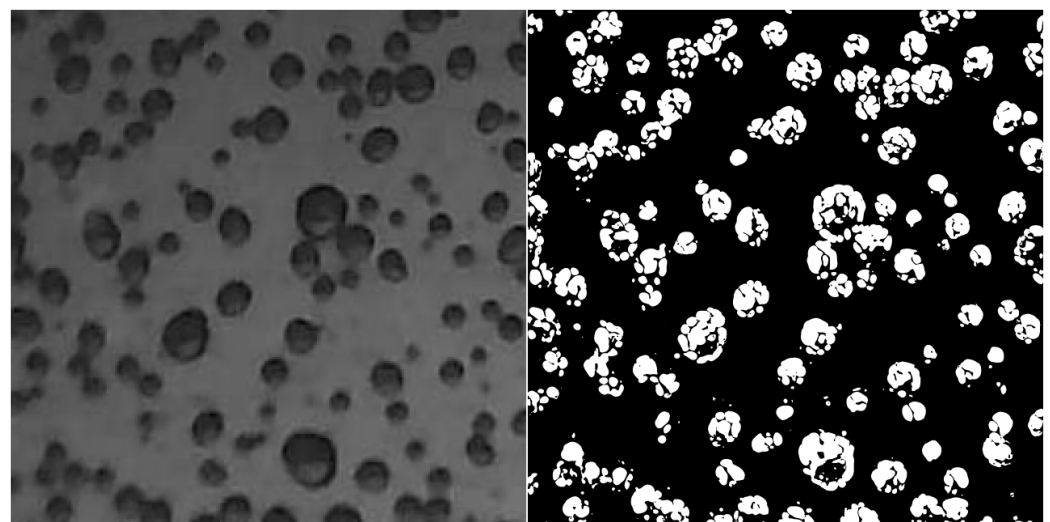
Figure 22. The segmentation effect of the model when the droplet size is too large.

3.2.2. Computing Power Requirements for the DSRSS Model

The number of parameters in the DSRSS model is less than 1 M. Since it is a fully convolutional network, it can accept inputs of different sizes. Since the encoder part of the model contains three times of downsampling, the length and width of the input image are required to be an integral multiple of eight. The computational complexity and RAM requirements of the model are positively correlated with the input size. Tests were

conducted on several computer hardware and the RAM capacity consumed by inferring single images of different sizes was recorded, the amount of computation, and the inference speed on different computing hardware. The results are shown in Table 2. Since the spray quality evaluation work does not require absolute real-time performance, its inference speed is tolerable on typical devices.

Table 2. Performance of models on different hardware platforms under different input sizes.

Input Size (Pixels)	RAM Consumption (GiB)	Calculation Quantity (GFLOPS)	Inference Speed (FPS)			
			Nvidia Tesla P40	Nvidia GTX 1660Ti Max-Q	Intel Core i7-9750H	Intel Xeon E3-1245V5
64 × 64	0.4	0.99	51.3	46.1	28.6	24.3
128 × 128	0.5	3.95	49.8	40.4	12.5	11.5
256 × 256	0.7	15.8	42.3	37.2	4.16	3.86
384 × 384	1.0	35.6	34.9	26.0	2.03	1.88
512 × 512	1.3	63.2	24.8	16.3	1.12	1.08

4. Conclusions and Future Work

The paper provides a lightweight end-to-end DSRSS model and a MACS algorithm to improve the accuracy of semantic segmentation in droplet deposition images and achieve precise and efficient multi-parameter measurement of droplet deposition with low-quality droplet deposition images for a scalable low-cost spraying measurement system for use in field scenarios. Based on DSRSS, MACS, and droplet area correction, a complete image processing procedure for droplet deposition images in a low-cost spray measurement system has been established. The main conclusions are as follows.

(1) When the two weight parameters α and β in *WIoU* are 0.775 and 0.225, respectively, the rate of independent droplet segmentation when using the DSRSS on the DDID reaches 0.998 and the *IoU* reaches 0.973.

(2) The MACS algorithm can reduce the adhesion rate of droplet deposition images with a coverage exceeding 30% by 15.7%, and the area correction function can reduce the area segmentation error in the DSRSS by 3.54%.

(3) The quantity of parameters in the DSRSS is less than 1 M, and it can reduce RAM requirements through image input blocking, thus being able run on embedded platforms.

(4) Based on the exact test metrics obtained from the test dataset and the visualization of segmentation in real droplet deposition images, the comprehensive performance of this method is much better than that of the adaptive threshold segmentation method, as its droplet boundary segmentation is more accurate and the accuracy of droplet area estimation is higher.

This work can be used as a solution for high-speed in situ spray measurement, and provides a reference for the segmentation of low-quality dense object images. The complexity of the in situ measurement environment, device aging, and wear, especially the aging of the light source and the wear of the droplet collector, as well as the residue on the droplet collector after long-term use, will all affect the actual effectiveness of the method. As the next step, we plan to further improve the robustness of the method and make it suitable for more complex application scenarios.

Author Contributions: J.L.: Writing—original draft, Methodology, Data curation, Validation, Software. S.Y.: Investigation, Validation. X.L.: Writing—Review and Editing, Conceptualization, Methodology, Funding acquisition. G.L.: Data curation, Validation. Z.X.: Methodology, Investigation, Validation. J.Y.: Investigation, Resources, Project administration, Supervision. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Anyone can access the data by sending an email to jyuan@sdau.edu.cn.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Makwana, A.; Mohnot, P. Effect of spraying using sprayer robot for cotton crop: Sprayer robot for cotton crop. *J. AgriSearch* **2022**, *9*, 255–259. [CrossRef]
2. Guo, N.; Liu, S.; Xu, H.; Tian, S.; Li, T. Improvement on image detection algorithm of droplets deposition characteristics. *Trans. Chin. Soc. Agric. Eng.* **2018**, *34*, 176–182. [CrossRef]
3. Sarghini, F.; Visacki, V.; Sedlar, A.; Crimaldi, M.; Cristiano, V.; de Vivo, A. First measurements of spray deposition obtained from UAV spray application technique. In Proceedings of the 2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Portici, Italy, 24–26 October 2019; pp. 58–61. [CrossRef]
4. Wang, P.; Yu, W.; Ou, M.; Gong, C.; Jia, W. Monitoring of the pesticide droplet deposition with a novel capacitance sensor. *Sensors* **2019**, *19*, 537. [CrossRef] [PubMed]
5. Wen, Y.; Zhang, R.; Chen, L.; Huang, Y.; Yi, T.; Xu, G.; Hewitt, A.J. A new spray deposition pattern measurement system based on spectral analysis of a fluorescent tracer. *Comput. Electron. Agric.* **2019**, *160*, 14–22. [CrossRef]
6. Yang, W.; Li, X.; Li, M.; Hao, Z. Droplet deposition characteristics detection method based on deep learning. *Comput. Electron. Agric.* **2022**, *198*, 107038. [CrossRef]
7. Zhu, H.; Salyani, M.; Fox, R.D. A portable scanning system for evaluation of spray deposit distribution. *Comput. Electron. Agric.* **2011**, *76*, 38–43. [CrossRef]
8. Wang, L.; Song, W.; Lan, Y.; Wang, H.; Yue, X.; Yin, X.; Tang, Y. A Smart Droplet Detection Approach With Vision Sensing Technique for Agricultural Aviation Application. *IEEE Sens. J.* **2021**, *21*, 17508–17516. [CrossRef]
9. Vacalebre, M.; Frison, R.; Corsaro, C.; Neri, F.; Santoro, A.; Conoci, S.; Anastasi, E.; Curatolo, M.C.; Fazio, E. Current State of the Art and Next Generation of Materials for a Customized Intraocular Lens according to a Patient-Specific Eye Power. *Polymers* **2023**, *15*, 1590. [CrossRef]
10. Gennes, P.-G.; Brochard-Wyart, F.; Quéré, D. *Capillarity and Wetting Phenomena: Drops, Bubbles, Pearls, Waves*; Springer: Berlin/Heidelberg, Germany, 2004.
11. Kupyn, O.; Budzan, V.; Mykhailych, M.; Mishkin, D.; Matas, J. Deblurgan: Blind motion deblurring using conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018. [CrossRef]
12. Kupyn, O.; Martyniuk, T.; Wu, J.; Wang, Z. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.
13. Quan, Y.; Wu, Z.; Ji, H. Gaussian kernel mixture network for single image defocus deblurring. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 20812–20824. [CrossRef]
14. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017. [CrossRef]
15. Wang, L.; Li, D.; Zhu, Y.; Tian, L.; Shan, Y. Dual super-resolution learning for semantic segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020.
16. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Change Loy, C. Esrgan: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) Workshops, Munich, Germany, 8–14 September 2018. [CrossRef]
17. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015. [CrossRef]
18. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
19. Chen, L.-C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018. [CrossRef]
20. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015. [CrossRef]
21. Ramachandran, P.; Zoph, B.; Le, Q.V. Searching for Activation Functions. *arXiv* **2017**, arXiv:1710.05941. [CrossRef]
22. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. Yolact: Real-time instance segmentation. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.

23. Terven, J.; Cordova-Esparza, D. A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond. *arXiv* **2023**, arXiv:2304.00501. [CrossRef]
24. Zhao, X.; Ding, W.; An, Y.; Du, Y.; Yu, T.; Li, M.; Wang, J. Fast Segment Anything. *arXiv* **2023**, arXiv:2306.12156. [CrossRef]
25. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y.; et al. Segment Anything. *arXiv* **2023**, arXiv:2304.02643.