

Article

Real-Time Detection of Electrohydrodynamic Atomization Modes via a YOLOv8-Based Deep Learning Model

Xiong Ran ¹, Heming Xu ², Xiangfei Wei ², Jinxin Wang ^{2,*} and Wei-Cheng Yan ^{1,*}¹ School of Chemistry and Chemical Engineering, Jiangsu University, Zhenjiang 212013, China² Department of Electronic and Information Engineering, Bozhou University, Bozhou 236800, China

* Correspondence: bzuu_wangjx@163.com (J.W.); yanwc@ujs.edu.cn (W.-C.Y.)

Abstract

A YOLOv8-based deep learning model was developed to address real-time detection and dynamic regulation needs of the electrohydrodynamic atomization process. An EHDA experimental system was built to obtain images of six typical atomization modes, forming a dataset with 6000 images. After annotation and mosaic augmentation, the dataset served as the training data for the model. The YOLOv8 adopts a “backbone-neck-head” architecture to extract and fuse features, decouple classification and detection, and optimize performance. Experimental results demonstrate that on the test set, the model attains a precision value, recall rate, and mAP50 of 0.995, alongside an mAP50-95 of 0.8. Additionally, its prediction accuracy exceeds 0.99 across all operational modes. Compared with 10 models, it has the best precision and mAP50, as well as low computational complexity, combining high accuracy and lightweight advantages, which can be effectively used for real-time detection of EHDA modes.

Keywords: electrohydrodynamic atomization (EHDA); YOLOv8; mode recognition; real-time detection

1. Introduction

Electrohydrodynamic atomization (EHDA) leverages electrostatic forces to deform conductive liquids into a Taylor cone-jet, which breaks into uniform micro/nanodroplets, making it an important technology for high-precision micro/nanomaterial preparation [1–3]. Due to its advantages such as simple device structure, convenient operation, high preparation precision, and good reproducibility, this technology demonstrates broad application potential in the preparation of micro- and nanomaterials in fields such as materials science [4–7], chemical engineering [8,9], biomedical engineering [10–13], and food engineering [14,15].

The EHDA process involves complex multi-physics coupling (electrostatic force, surface tension, viscous force), and its atomization mode is highly sensitive to operating parameters (voltage, flow rate, nozzle structure) and liquid properties (conductivity, dielectric constant). Six typical modes have been identified: atomization, cone-jet, dripping, spindle, rotational, and skewed jet-atomization [16]. Among these modes, the stable atomization mode can form a steady jet at the apex of the Taylor cone and produce highly monodispersed uniform droplets [17,18], demonstrating significant advantages in particle preparation. Electrohydrodynamic atomization necessitates the application of ultra-high-voltage electrostatic fields to disintegrate charged media into micro/nanoscale droplets.



Academic Editor: Dewei Wang

Received: 13 December 2025

Revised: 6 January 2026

Accepted: 10 January 2026

Published: 15 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.

This inherently hazardous process demands real-time safety monitoring mechanisms. Concurrently, EHDA exhibits extreme dynamism and parametric sensitivity, with operational states undergoing significant transitions within millisecond timescales—rendering manual intervention ineffective for precision control. These dual characteristics necessitate the closed-loop integration of pattern recognition with adaptive control systems to maintain process stability.

Over the past few years, with the rapid advancement of AI technology, Machine Learning (ML) methods have been extensively applied to the mode recognition of EHDA to achieve efficient and accurate mode recognition and parameter optimization [19–25]. For instance, Kim et al. [22] utilized a 1-D CNN (One-Dimensional Convolutional Neural Network) model to extract current signals during the electrohydrodynamic spraying process and, through training, classified them with a model achieving an accuracy rate of 96.30%. Furthermore, Sun et al. [23] employed 4000 electrohydrodynamic printing images to train a Convolutional Neural Network (CNN) model, successfully classifying 8 modes, thereby developing an effective detection system to rapidly obtain the optimal printing mode. These ML model-based recognitions modes are of great significance as they can substantially reduce the workload.

However, the aforementioned traditional CNN-based methods essentially fall under the category of image classification tasks and can only output pattern categories, failing to achieve the localization of spray regions. Additionally, they require additional preprocessing steps (such as signal segmentation and image cropping) to interface with dynamic control, resulting in detection delays that make it difficult to meet the real-time regulation requirements of EHDA. Therefore, developing specialized real-time monitoring systems is not only a prerequisite for implementing intelligent closed-loop control, but also a key technical requirement for ensuring industrial safety production and improving process precision.

YOLO (You Only Look Once) as a detection algorithm, reframes detection as a regression problem and enables end-to-end real-time prediction [26–29]. YOLOv8, introduced by Ultralytics (Madrid, Spain) in January 2023, has been extensively utilized across a wide array of fields, due to its exceptional processing speed and real-time performance [30]. For instance, Duan et al. [31] utilized YOLOv8 to detect the pecking behavior of Cherry Valley ducks in real time, achieving a detection speed of 76.3 frames per second (FPS) and an mAP of 90.24%, demonstrating outstanding detection accuracy. Similarly, Wang et al. [32] employed YOLOv8 to monitor the use of safety equipment by workers in complex mining environments, effectively enhancing operational safety for miners. Additionally, Guo et al. [33] improved the YOLOv8s model to address real-time object detection challenges in underwater complex environments, achieving mAP scores of 84.3%, 84.7%, and 86.8% on three custom datasets, with a detection speed reaching 156 frames per second (FPS). These applications highlight YOLOv8's exceptional detection performance and strong real-time capabilities.

YOLOv8, as an end-to-end object detection algorithm, has the core advantage of simultaneously accomplishing “spray region localization and pattern classification” in a single inference pass, without requiring any intermediate preprocessing steps. It may reduce detection latency to the millisecond level, thereby precisely aligning with the timing requirements of the “pattern recognition–parameter adjustment” closed-loop control in the EHDA process. Based on this, this paper innovatively applies the YOLOv8 object detection framework to electrohydrodynamic atomization pattern recognition tasks, achieving a strategically significant technological transition: a paradigm shift from post-event analysis to continuous real-time detection. This breakthrough not only effectively addresses the core limitations of traditional machine learning methods—“classification-only capability, lack of localization, and insufficient real-time performance”—but also enables a fundamental

optimization tailored to the inherent spatiotemporal evolution characteristics of EHDA processes. This technological transition represents more than a simple model adaptation; it establishes the technical foundation for real-time monitoring in smart closed-loop control systems within industrial EHDA scenarios, directly advancing the critical leap toward intelligent manufacturing transformation. The specific work includes the following three aspects: (1) Constructing an annotated dataset containing six EHDA patterns for training the YOLOv8 model; (2) Validating the model's ability to simultaneously achieve atomization region localization and pattern classification, and systematically evaluating its real-time performance (detection speed, latency); (3) Comparing the proposed model with 10 mainstream detection models to validate its superior performance in EHDA pattern detection tasks.

2. Experiments and Model Development

2.1. EHDA Experimental System for Real-Time Detection

To capture dynamic EHDA images and integrate real-time detection, an experimental system was built as shown in Figure 1, consisting of three core modules: (1) EHDA generation module, (2) high-speed image acquisition module, and (3) YOLOv8 real-time detection module. During the experiment, the nozzle is first connected via a conduit to the outlet of the syringe pump to ensure that the liquid can flow stably from the nozzle at the set flow rate. Then, the positive electrode of the high-voltage power supply is connected to the nozzle, and the negative electrode is connected to the grounded substrate; this setup enables the establishment of a high-voltage electric field between the nozzle and the substrate with a fixed spacing. The electric field is used to accelerate the flow of liquid from the nozzle tip and promote the formation of various modes. Throughout the experiment, a high-speed camera captures and records images of the EHDA process. The shooting frame rate is set at 20 FPS, and an LED (Light Emitting Diode) light is used for background fill lighting to enhance image clarity. This study adopted a 20 FPS frame rate (based on reference [21]), which sufficiently captures EHDA mode transitions occurring at tens of milliseconds timescale. This sampling frequency (50 ms/frame) effectively records morphological differences between modes while balancing data acquisition efficiency and storage requirements.

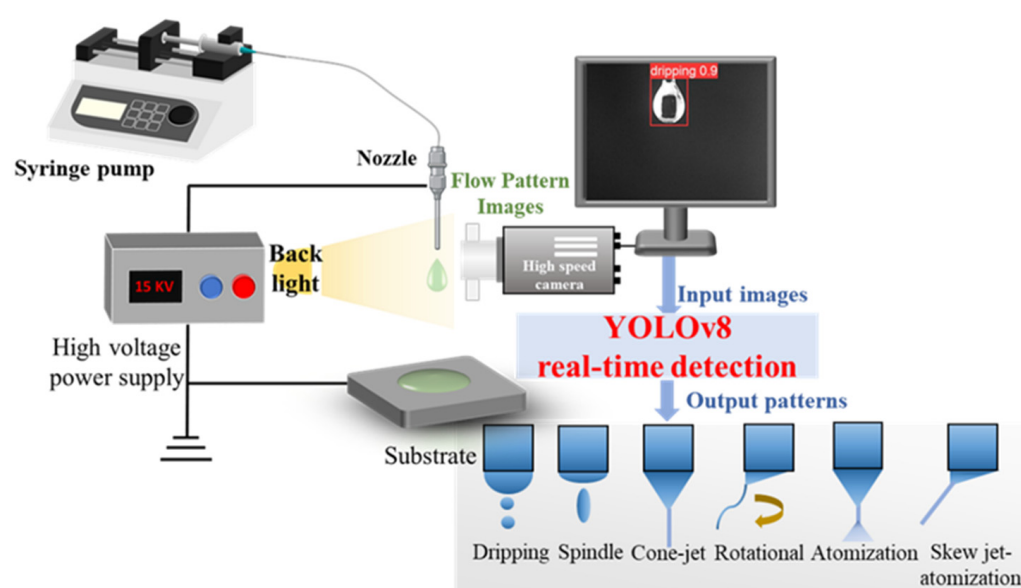


Figure 1. Schematic of EHDA experimental system coupled with mode detection process.

By adjusting the solution properties, operational parameters, and nozzle structure, different atomization modes could be achieved. To ensure the generalizability of the dataset, the conductivity of the solutions used in the experiment ranged from 2 to 10 $\mu\text{S}/\text{cm}$, the viscosity ranged from 20 to 100 mPa·s, the applied voltage ranged from 1 to 30 kV, the liquid volume flow rate ranged from 5 to 30 mL/h, the distance from the nozzle to the substrate ranged from 2 to 10 cm, and five different nozzle sizes were employed. A dataset comprising 6000 jet images was established, with images categorized into six typical electrohydrodynamic atomization regimes based on the majority voting principle: dripping mode, spindle mode, cone-jet mode, whipping mode, atomization mode, and skewed jet-atomization mode. This classification comprehensively covers the commonly observed atomization morphologies in engineering applications. Detailed results of the mode recognition and associated discussions have been presented in our previous work and are thus referred to herein for further reference [21].

2.2. Data Preprocessing

In electrohydrodynamic atomization pattern recognition, discrimination primarily depends on morphological features of fluid interfaces—such as conical structures, jet stability, and droplet distribution patterns—rather than chromatic information. Grayscale imaging is therefore employed to preserve these critical structural characteristics while substantially reducing data dimensionality and computational complexity. The image dataset captured from the EHDA experiments was annotated using the LabelImg (v1.8.6) tool, where targets in the images were bounding-box labeled with their respective categories, generating corresponding annotation files. Each annotation file consists of the labeled category, the coordinates (x, y) of the target center point, and the actual proportional dimensions (width, height) of the bounding box relative to the entire image. Figure 2a shows the input quantities of the six mode categories. In this study, 1000 images per category were collected to construct the dataset, which was divided into training, and test sets in an 8:2 ratio.

Figure 2b illustrates the model's focus areas for object detection (or anchor box matching). The red/highlighted regions in the figure represent the core areas where the model identifies targets, used to evaluate the alignment between anchor boxes and the actual target shapes, a few light-yellow boxes indicate areas of attention dispersion. This visualization demonstrates whether the anchor box designs conform to target morphologies (e.g., whether elongated targets are matched with appropriately shaped elongated anchor boxes) and whether the model can consistently focus on key parts of targets (such as centers or edges). The figure shows that the model primarily employs square, elongated, and rectangular anchor boxes, a distribution fully consistent with the actual target morphologies—including dripping, spindle, and cone-jet modes, as well as atomization, skew cone-atomization, rotational modes, and dispersed droplets at the tail of the spray. The results indicate that the model's core recognition areas for targets highly align with their actual morphologies in this dataset.

Figure 2c,d are feature scatter plots of the dataset, illustrating the distribution modes of target features (coordinates and dimensions). According to Figure 2c, the color tends to darken around $x = 0.482$ and $y = 0.335$, indicating a higher density distribution. This suggests that targets frequently appear in the upper-central region of the image. Based on Figure 2d, an analysis of target size correlations leads to the conclusion that the data predominantly clusters in three linear band-like distributions: width = 0.73, height = 0.48 and height = 2 width = 0.1. This pattern indicates a predictable width-to-height ratio across targets of different sizes, and the model leverages this data characteristic to assist in anchor box optimization.

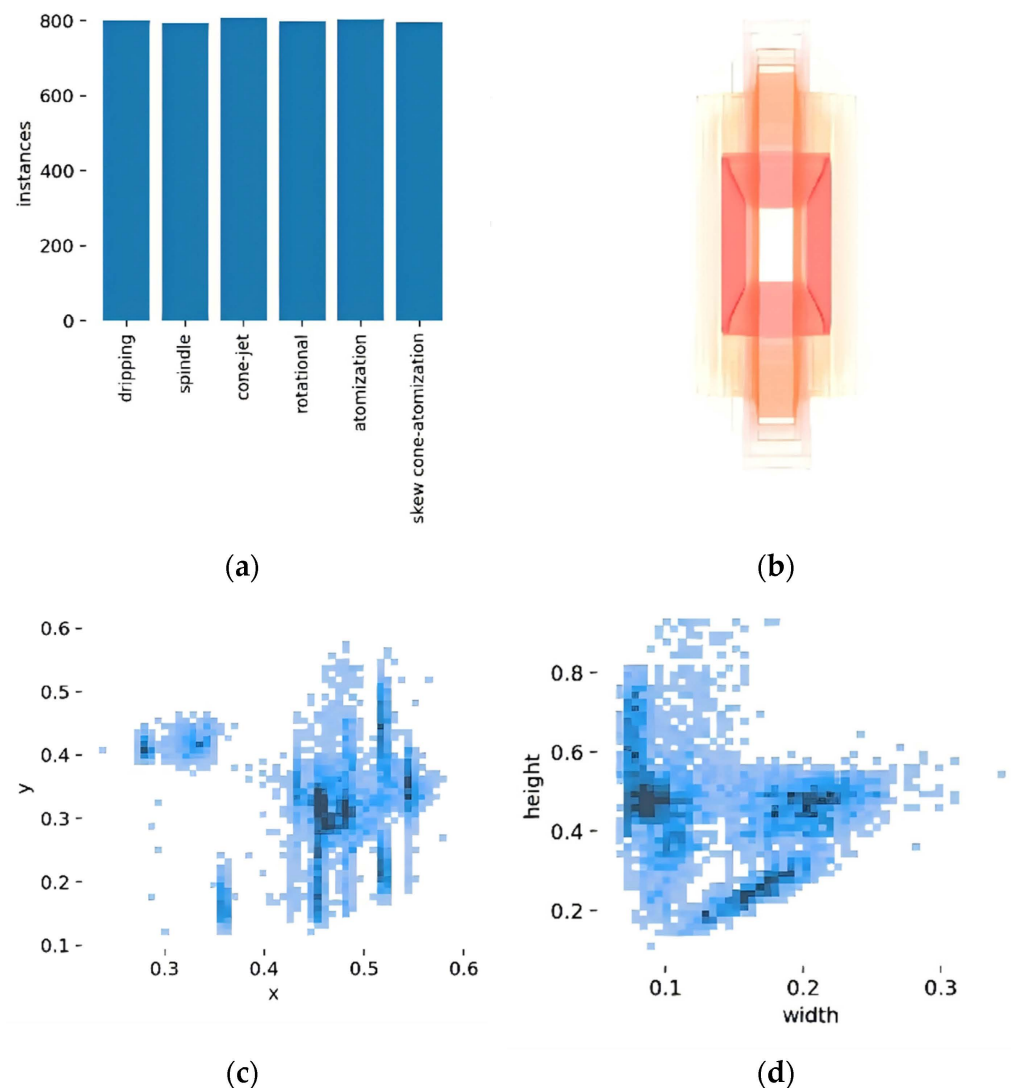


Figure 2. Input feature map: (a) Category and quantity of input images (b) Anchor box design diagram (c) Center point scatter plot (d) Width-Height scatter plot.

Although the nozzle position and angle variations during the experiment were limited, to further enhance the model's generalization capability under different imaging conditions (such as target scale, position, and background combinations), we employed the Mosaic data augmentation technique. This method involves randomly selecting four training images for cropping and stitching, simulating potential positional shifts, scale variations, and multi-modal overlaps of targets in the images, thereby expanding the visual diversity of the training data, as specifically illustrated in Figure 3. Although this approach does not cover all possible extreme imaging scenarios, the high accuracy and strong robustness demonstrated on the validation and test sets indicate that Mosaic augmentation is sufficient to effectively support the model's generalization requirements in practical EHDA detection tasks. Each synthesized image is a grayscale image, enabling us to learn more features with a smaller-scale training set. Figure 3 also shows that during the data annotation process, all six spray categories were clearly marked with bounding boxes, indicating high annotation accuracy with no missed instances.

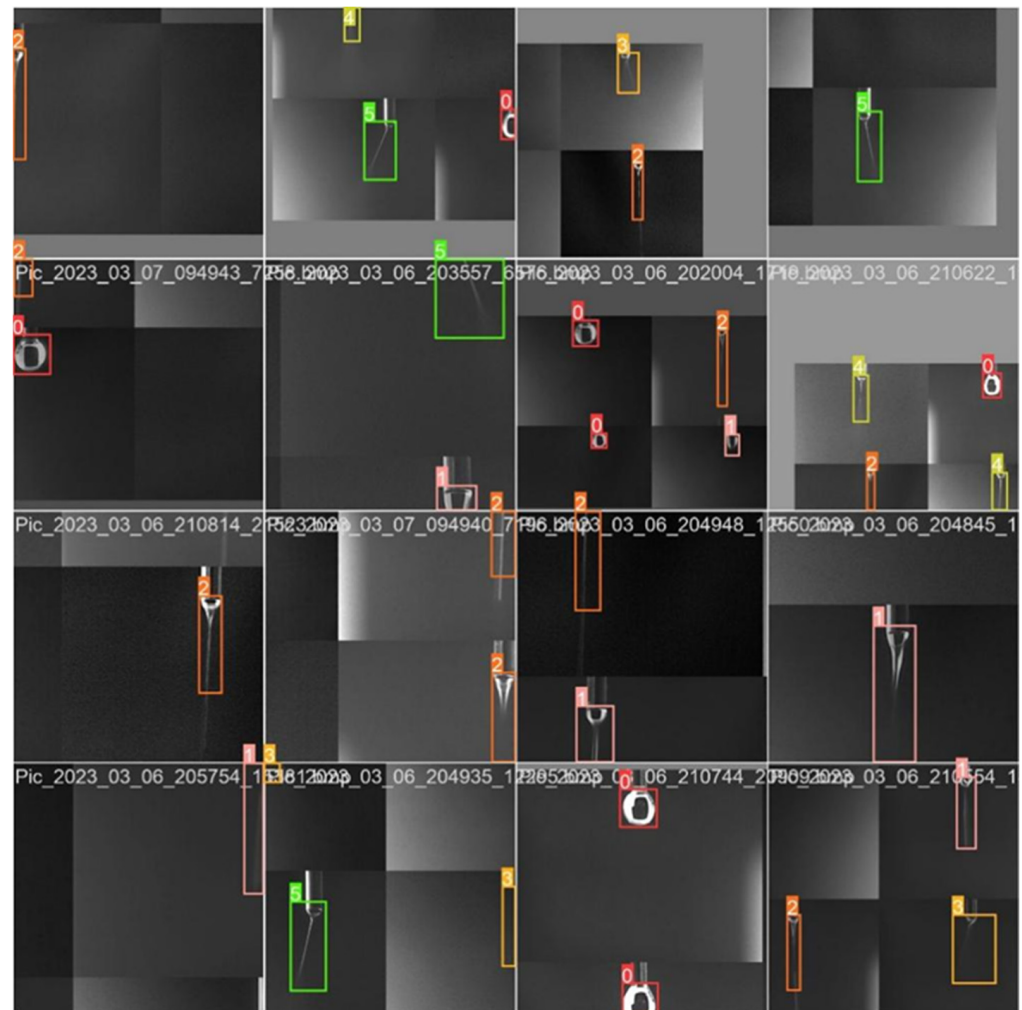


Figure 3. Mosaic-augmented images of EHDA modes with bounding box annotations.

2.3. YOLOv8 Model

YOLOv8 achieves an optimal trade-off between detection accuracy and computational efficiency. Its unique architecture exhibits strong alignment with the spatiotemporal characteristics of electrohydrodynamic atomization processes, enabling precise real-time tracking of fluid trajectories and morphological evolution. Crucially, YOLOv8 demonstrates empirically validated stability and maturity compared to later YOLO variants—attributes essential for meeting the rigorous reliability demands of industrial deployment. Consequently, it was selected as the foundational model for EHDA image recognition tasks. As shown in Figure 4, the YOLOv8 network architecture is composed of three parts: backbone, neck, and head. First, the Backbone extracts multi-level and multi-scale feature information from the original image, providing the foundation for subsequent detection and classification. Then, the Neck integrates the feature information extracted by the Backbone at different levels and scales, constructing a feature pyramid that incorporates rich spatial details (from shallow layers) and high-level semantic information (from deep layers), thereby improving detection accuracy. Finally, the Head receives the feature pyramid (multi-scale feature maps) from the Neck and predicts the category and location of targets at each pixel of every feature map.

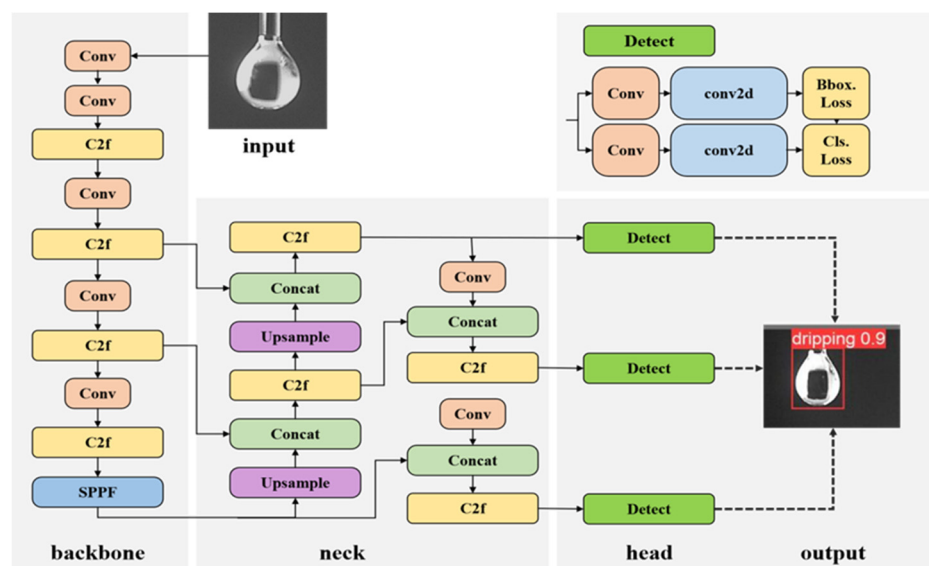


Figure 4. YOLOv8 network architecture diagram.

The backbone network comprises convolutional (Conv), Conv-to-Fully-Connected (C2f), and Spatial Pyramid Pooling Fusion (SPPF) structures. As shown in Figure 5a, the C2f module processes the input feature map by first performing a convolution operation, then splitting the output into two parts: one part is directly propagated forward, while the other undergoes further processing through multiple Bottleneck structures (composed of several Conv layers). Finally, all results are concatenated. This Split Add Concatenation structural design not only strengthens gradient flow but also achieves effective integration of shallow detail features (such as droplet edges and jet textures) with deep semantic features (such as the overall contour of the pattern). It is particularly suitable for identifying diverse modes in EHDA (e.g., the slender linear structures of cone-jet, the dispersed particulate structures of atomization, and the helical structures of rotational modes), addressing the challenge of distinguishing similar patterns using single-scale features. Meanwhile, the SPPF module expands the receptive field significantly with low computational cost by serially stacking pooling layers, enabling the model to simultaneously capture local details and global contextual information. This effectively handles the substantial variation in target scales in EHDA (from droplets of hundreds of micrometers in dripping mode to particles of only a few micrometers in atomization mode). The synergistic effect of these two modules significantly improves detection accuracy and robustness for multi-scale, multi-form EHDA targets while maintaining lightweight characteristics and low latency.

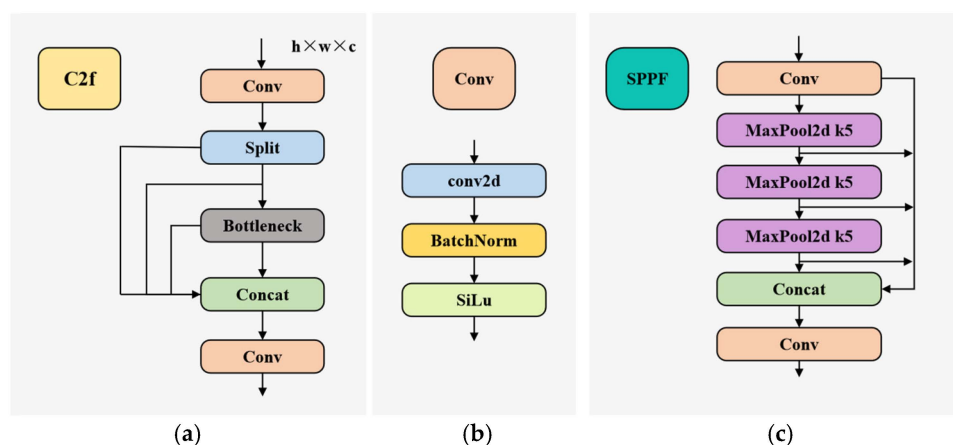


Figure 5. Network structure diagram: (a) C2f; (b) Conv; (c) SPPF.

The Neck acts as a neural hub connecting the Backbone and the Head. By integrating the Feature Pyramid Network (FPN) and the Path Aggregation Network (PAN), it establishes a bidirectional feature fusion architecture, enabling multi-scale detection capabilities while balancing semantic information and spatial precision.

In YOLOv8, the Head component decouples the processes of object classification and bounding box detection. This design primarily involves two key steps: first, loss computation, and second, the filtering of detected bounding boxes. The most critical aspect of the loss computation process is the design of the positive-negative sample assignment strategy and the loss function. YOLOv8 mainly adopts the Task Aligned Assigner strategy, which evaluates sample quality using a measurable metric calculated by the following formula: $alignment_metric = s^\alpha \times u^\beta$. This formula combines classification confidence (s) with the Intersection over Union (IoU) between predicted and ground truth bounding boxes (u), thereby balancing the selection of anchor samples that best match the task objectives.

YOLOv8 employs the Complete IoU (CIoU) loss to calculate the bounding box regression loss. For a predicted box $b = (x, y, w, h)$ and its corresponding ground truth box $b_{gt} = (x_{gt}, y_{gt}, w_{gt}, h_{gt})$, the CIoU loss is defined as:

$$\mathcal{L}_{CIoU} = 1 - IoU + \frac{\rho^2(b, b_{gt})}{c^2} + \alpha v \quad (1)$$

The variables in Equation (1) are defined as follows: ρ denotes the Euclidean distance between the centers of the two boxes, calculated as $\rho = \sqrt{(x - x_{gt})^2 + (y - y_{gt})^2}$; c is the diagonal length of the smallest enclosing box that covers both the predicted and ground truth boxes; v measures the discrepancy in aspect ratio, defined as $v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w}{h} \right)$; and α is a balancing parameter given by $\alpha = \frac{v}{1 - IoU + v}$, which adjusts the relative importance of different geometric deviations in the loss function.

The Classification Loss in YOLOv8 is calculated using BCE Loss (Binary Cross-Entropy Loss), with the formula given by:

$$\mathcal{L}_{cls} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C [y_{i,c} \cdot \log(\sigma(\hat{y}_{i,c})) + (1 - y_{i,c}) \cdot \log(1 - \sigma(\hat{y}_{i,c}))] \quad (2)$$

where y denotes the ground truth label, $\hat{y}$ represents the raw output logits from the model, and σ is the Sigmoid activation function, which maps the logits to predicted probabilities in the range (0, 1). The parameter N indicates the number of samples in the batch (typically including both positive and negative examples), and c corresponds to the total number of classes.

DFL (Distribution Focal Loss) is used for the regression prediction of bounding box coordinates. Its mathematical expression is:

$$\mathcal{L}_{DFL}(S_i, S_{i+1}) = -((y_{i+1} - y) \log(S_i) + (y - y_i) \log(S_{i+1})) \quad (3)$$

where y denotes the continuous ground truth value, y_i and y_{i+1} represent the two closest discrete anchor points to y , S_i and S_{i+1} are the Softmax-normalized probabilities predicted by the model for y_i and y_{i+1} , respectively.

2.4. Performance Metrics

Precision refers to the proportion of actual positive samples among those predicted as positive, which reflects the model's "accuracy capability". Higher precision indicates fewer false alarms (false positives) and more reliable prediction results. It is calculated by

dividing the number of true positives (TP) by the sum of true positives and false positives (TP + FP), where FP represents false positives.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

Recall is the proportion of actual positive instances that are correctly predicted by the model among all true positive cases. A higher recall value indicates fewer misses (false negatives) and stronger comprehensiveness in detection. It is calculated as the ratio of true positives to the total number of actual positive instances (TP + FN), where FN represents false negatives.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

AP (Average Precision) is a comprehensive metric that evaluates a model's detection capability for a single category by calculating the area under the precision-recall curve across all confidence thresholds. A higher AP value indicates more accurate and complete recognition for that category. mAP is a widely used metric designed to evaluate the accuracy and effectiveness of object detection models. Its calculation involves averaging the AP values for each object category present in the dataset. mAP50 (also known as mAP@0.5) computes the average AP across all categories with the IoU threshold fixed at 0.5, primarily evaluating the model's overall detection performance under "relatively lenient localization" conditions. In contrast, mAP50-95 (also known as mAP@0.5:0.95) is a stricter metric. It calculates the average mAP across 10 different IoU thresholds ranging from 0.5 to 0.95 (with a step of 0.05), which comprehensively reflects the model's overall performance from "coarse localization" to "precise bounding box selection."

$$AP = \int_0^1 P(R) dR \approx \frac{1}{N} \sum_{k=1}^N P_{\text{interp}}(R_k) \quad (6)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (7)$$

$$mAP_{50} = \frac{1}{C} \sum_{i=1}^C AP_i^{\text{IoU}=0.5} \quad (8)$$

$$mAP_{50:95} = \frac{1}{10} \sum_{t=50}^{95} AP_t \quad (9)$$

3. Results

3.1. Training Process

This study employs the Adam optimizer for parameter optimization (learning rate of 0.01, weight decay of 0.005). The training process was set with a maximum of 1000 iteration epochs, and an early stopping mechanism was introduced: training automatically terminates when the performance metrics on the validation set show no improvement over 50 consecutive training epochs. The evolution trends of various loss functions and performance metrics during training are shown in Figure 6, where the loss functions demonstrated stable convergence, while key performance metrics (including precision, recall, and average precision) consistently exhibited gradual optimization trends. It is noteworthy that the model reached the performance convergence threshold at approximately 300 training epochs.

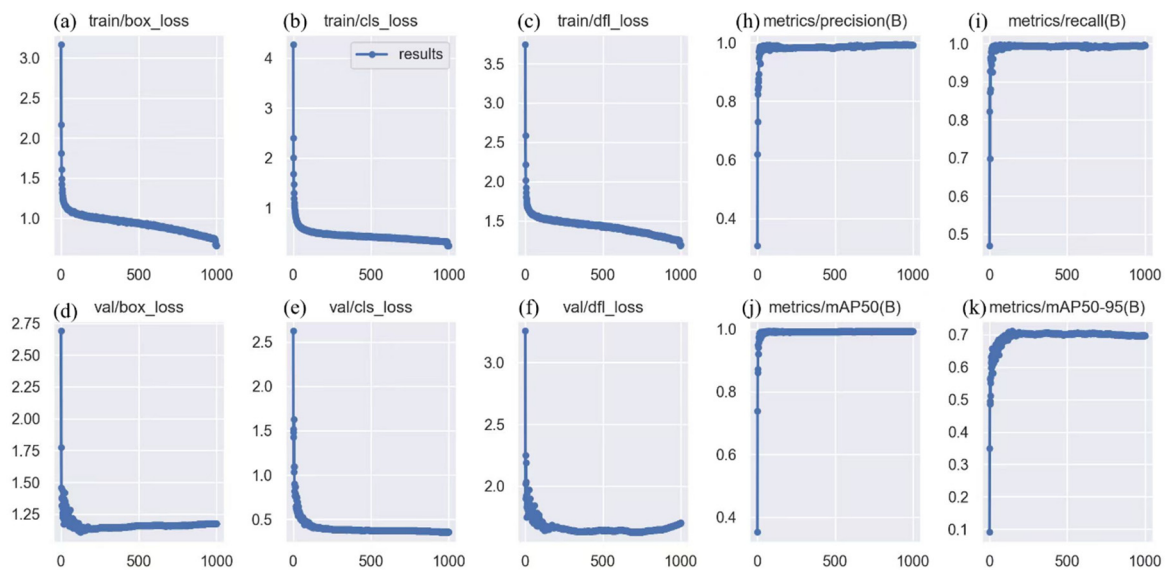


Figure 6. YOLOv8 training curves: (a) Training box loss, (b) Training cls loss, (c) Training DFL loss, (d) Validation box loss, (e) Validation cls_loss, (f) Validation DFL loss, (h) Validation Precision, (i) Validation Recall, (j) Validation mAP50, (k) Validation mAP50-95.

Figures 6a and 6d illustrate the changes in bounding box loss (box_loss) on the training and validation sets, respectively. As the number of training epochs increased, the bounding box loss on the training set gradually decreased, eventually stabilizing at approximately 0.5; the corresponding loss on the validation set also showed a declining trend, converging to around 1.0, reflecting the model's strong generalization capability in bounding box regression tasks. The classification loss (cls_loss) on the training and validation sets is shown in Figures 6b and 6e, respectively. Both decreased progressively during training, eventually stabilizing at around 1.0 and 2.0, indicating continuous improvement in the model's performance for object classification tasks. YOLOv8 introduces Distribution Focal Loss to optimize bounding box regression accuracy, with its changes on the training and validation sets shown in Figure 6c,f. This loss consistently decreased and eventually converged to approximately 1.5, demonstrating gradual improvement in the model's performance in this aspect. On the validation set, the model's Precision steadily increased during training (Figure 6h), eventually exceeding 1.0, reflecting its growing reliability in high-confidence detections. The Recall also showed an upward trend (Figure 6i), eventually reaching approximately 1.0, indicating the model's continuously enhanced ability to identify positive samples. Additionally, the mean Average Precision at an IoU threshold of 0.5 (mAP50) gradually improved and stabilized at around 1.0 (Figure 6j), highlighting the model's overall performance advantage in object detection tasks. The mAP50-95 metric also exhibited an upward trend (Figure 6k), eventually reaching approximately 0.8, demonstrating the model's robust detection capability under different IoU threshold conditions. In summary, all loss functions of YOLOv8 steadily converged during training, and performance metrics continued to improve, indicating excellent optimization stability and learning efficiency. The model achieved outstanding comprehensive performance on the validation set, with Precision, Recall, mAP50, and mAP50-95 approaching approximately 1.0, 1.0, 1.0, and 0.8, respectively, fully validating its competitiveness in object detection tasks. The distinct visual characteristics of EHDA modes, combined with YOLOv8's ability to capture multi-scale features, enable the model to achieve near-perfect metrics without overfitting, as validated by consistent performance on unseen test data.

3.2. Feature Attention Mechanism

Figure 7 presents the visualization results of heatmaps generated using the Grad-CAM (Gradient-weighted Class Activation Mapping) method to interpret the feature extraction mechanism of the YOLOv8 model during EHDA mode recognition. The heatmaps use color variations to represent the contribution levels of different regions in the image to the model's final decision, with red areas indicating the most significant impact on the recognition outcome, followed by yellow and orange regions, while blue areas denote the lowest contribution. Notably, regions with higher feature contributions are primarily distributed along the actual flow paths of the conductive liquid, and within these areas, locations with clearer morphological features and more typical structures exhibit correspondingly higher contributions—morphological characteristics being the core basis for discriminating EHDA modes. Unlike rigid targets (such as vehicles in object detection), the EHDA modes—particularly the atomization mode and the tilted jet–atomization mode—exhibit boundary blurring due to droplet scattering, where droplets disperse irregularly, and those in the marginal regions still carry critical physical features. The scattered attention observed in certain areas of the figure is not a flaw but a functional design: it reflects both the inherent irregularity of electrospray technology and ensures robust real-time monitoring performance. Overall, the visualization results in Figure 7 clearly demonstrate that the model's attention regions align precisely with morphologically critical areas in practical applications, indicating YOLOv8's effectiveness in capturing visual features most relevant to atomization mode discrimination. Thus, it can be concluded that the model constructed in this study possesses excellent feature extraction and identification capabilities, accurately refining key information to support reliable classification of atomization modes, further validating its effectiveness and interpretability in EHDA mode analysis tasks.

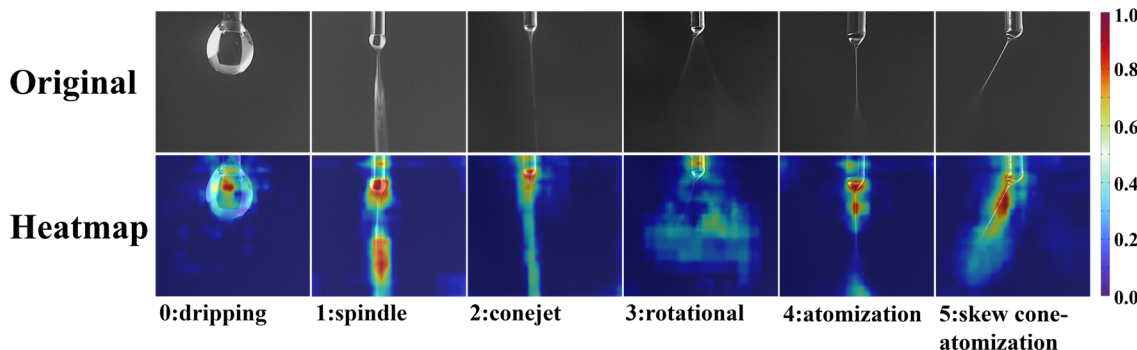


Figure 7. Attention Mechanism Heatmap for six EHDA modes.

3.3. Mode Detection Performance

Figure 8 presents the confusion matrix of the final YOLOv8 model's recognition results for different EHDA modes on the test set, which visually reflects the model's performance in the multi-category classification task. It can be clearly observed from the figure that the values on the diagonal are significantly higher than those in the off-diagonal regions, with the model maintaining a prediction accuracy above 0.99 for all categories. This result fully demonstrates the model's exceptionally high classification confidence and stability.

Furthermore, Figure 9 showcases examples of the model's detection performance on real images, revealing that the model not only accurately locates the target regions but also successfully identifies the corresponding EHD spray modes, exhibiting superior end-to-end discrimination capability. Based on these comprehensive results, it can be concluded that the constructed YOLOv8 model can be effectively applied to mode analysis and condition detection of EHDA processes.

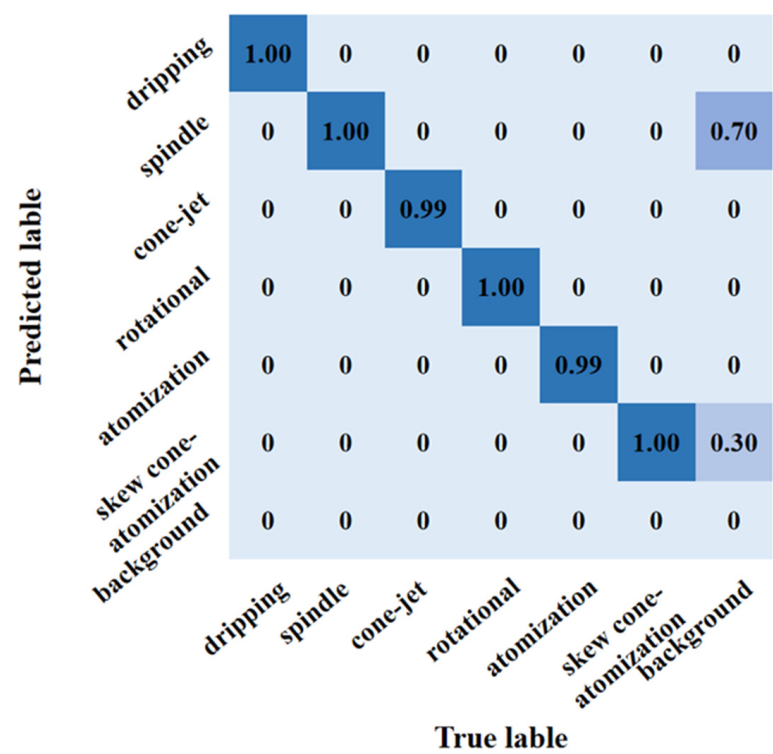


Figure 8. Confusion matrix chart.

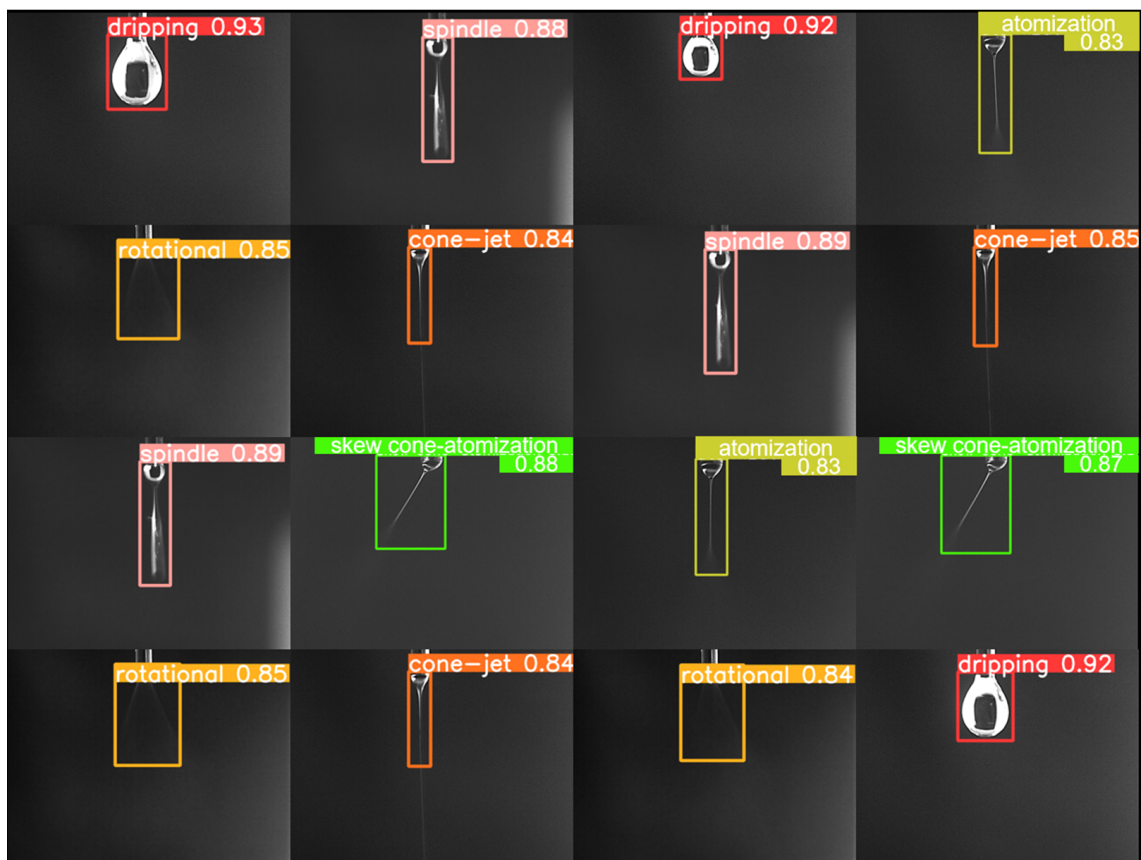


Figure 9. Detection results display diagram.

To validate the comprehensive performance of the YOLOv8 model proposed in this study, 11 detection models—Faster RCNN (Faster Regions with Convolutional Neural Networks), EfficientDet (Efficient Object Detection), SSD (Single Shot MultiBox Detec-

tor), DETR (DEtection TRansformer), YOLOv3, YOLOv5, YOLOv6, YOLOv9, YOLOv10, YOLOv11, and YOLOv8—were utilized for comparison. The performance test results are presented in Table 1.

Table 1. Models evaluation results.

Model	Precision	Recall	mAP@0.5	GFlops	Params (M)
Faster RCNN	0.971	0.986	0.989	30.6	13.3
EfficientDet	0.974	0.987	0.990	55	20.7
SSD	0.976	0.988	0.991	30.7	13.3
DETR	0.989	0.986	0.991	86	41
YOLOv3	0.987	0.986	0.991	154.7	61.5
YOLOv5	0.858	0.870	0.899	16.5	7.2
YOLOv6	0.991	0.989	0.993	42.8	15.9
YOLOv9	0.992	0.993	0.994	58.3	18.7
YOLOv10	0.987	0.996	0.994	24.5	8.03
YOLOv11	0.987	0.996	0.994	21.3	9.4
YOLOv8	0.995	0.995	0.995	23.4	9.8

Table 1 demonstrates the performance of various detection models in the EHDA mode detection task. From a comprehensive perspective, the YOLOv8 model proposed in this paper exhibits superior performance across most evaluation metrics. Specifically, YOLOv8 achieves 0.995 in both precision and mAP@0.5, significantly outperforming other comparative models. Its recall rate also reaches 0.995, ranking second only to YOLOv10 and YOLOv11, yet still exceeding all other models. Furthermore, YOLOv8 demonstrates significantly lower computational complexity—as measured by GFLOPs (Giga Floating-point Operations per Second)—and reduced parameters (Params) compared to benchmark models. This indicates substantially decreased computational resource demands and enhanced suitability for embedded deployment. These experimental results comprehensively demonstrate that YOLOv8 maintains exceptionally high detection accuracy and recall rates while achieving outstanding computational efficiency and lightweight architecture. This balance underscores its capability for efficient and precise object detection, highlighting significant practical application potential.

4. Conclusions

This study developed a YOLOv8-based deep learning model to enable real-time detection and dynamic regulation of the electrohydrodynamic atomization process, with comprehensive validation using experimental data and comparative tests. A 6000-image dataset covering six EHDA modes was generated via a customized experimental system. After LabelImg annotation and mosaic augmentation, the dataset provided reliable training data for the YOLOv8 model. The model not only achieved stable convergence (training box loss stabilized at ~0.5, validation box loss at ~1.0) but also attained Precision, Recall, and an mAP50 of ~0.995 on the validation set. Compared with 10 competing models (e.g., Faster RCNN, YOLOv3–v11), the model exhibited the highest Precision and mAP@0.5, and the model maintained greater than 0.99 accuracy for all modes. In addition, Grad-CAM-generated attention heatmaps of the YOLOv8 model directly confirm that the model effectively captures the critical morphological characteristics that distinguish EHDA modes, further validating its interpretability and reliability for mode classification tasks. Overall, the model developed in this study demonstrates high performance in real-time EHDA detection, validating that lightweight deep learning models can achieve industrial-grade real-time requirements while maintaining high accuracy. Future research will focus on the

following directions: (1) Expanding dataset scale by incorporating more diverse operational conditions, fluid properties, and environmental interference factors to enhance model robustness; (2) conducting field tests in industrial environments to verify system stability and adaptability under complex working conditions and address potential engineering challenges during actual deployment; (3) integrating a real-time feedback control system to directly convert pattern recognition results into control parameter adjustments, thereby establishing a complete closed-loop control architecture; (4) exploring edge computing and model lightweighting technologies to enable efficient operation on resource-constrained industrial edge devices; and (5) developing cross-fluid and cross-device transfer learning strategies to reduce model training costs in new application scenarios.

Author Contributions: Conceptualization, X.R. and J.W.; methodology, X.R.; validation, X.R., H.X. and X.W.; formal analysis, H.X.; investigation, X.W.; resources, W.-C.Y.; data curation, X.R.; writing—original draft preparation, X.R.; writing—review and editing, J.W. and W.-C.Y.; visualization, X.W.; supervision, W.-C.Y. funding acquisition, J.W. and W.-C.Y. All authors have read and agreed to the published version of the manuscript.

Funding: The authors thank the supports from National Key Research and Development Program of China (2024YFA1510003), Jiangsu Provincial Outstanding Youth Fund Project (BK20250053), Anhui Provincial Natural Science Foundation (2408085QE164), Scientific Research Key Project of Education Department of Anhui Province (2024AH051309), Project for Cultivating Academic Leaders (DTR2024053), and Bozhou University Research Launch Fund (BYKQ2021Z01, BYKQ202514).

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

Abbreviations

The following abbreviations are used in this manuscript:

CNN	Convolutional Neural Network
EHD	electrohydrodynamic
EHDA	electrohydrodynamic atomization
YOLO	You Only Look Once
LED	Light Emitting Diode
mAP	mean Average Precision
Conv	convolutional
C2f	Conv-to-Fully-Connected
SPPF	Spatial Pyramid Pooling Fusion
FPN	Feature Pyramid Network
PAN	Path Aggregation Network
IoU	Intersection over Union
CIoU	Complete Intersection over Union
AP	Average Precision
cls_loss	classification loss
DFL	Distribution Focal Loss
Grad-CAM	Gradient-weighted Class Activation Mapping
GFLOPs	Giga Floating-point Operations per Second
Params	parameters
1-D CNN	One-Dimensional Convolutional Neural Network
ML	Machine Learning

References

- Xie, J.; Jiang, J.; Davoodi, P.; Srinivasan, M.P.; Wang, C.-H. Electrohydrodynamic atomization: A two-decade effort to produce and process micro-/nanoparticulate materials. *Chem. Eng. Sci.* **2015**, *125*, 32–57. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xue, J.; Wang, Z.; Chen, Y.; Wang, J.; Tian, J.; Li, B.; Wang, X.; Xu, H.; Huo, Y.; Dong, Q.; et al. Cone-jet regime in electrospray: A comprehensive review. *Phys. Fluids* **2025**, *37*, 081307. [\[CrossRef\]](#)
- Chen, L.; Ru, C.; Zhang, H.; Zhang, Y.; Wang, H.; Hu, X.; Li, G. Progress in Electrohydrodynamic Atomization Preparation of Energetic Materials with Controlled Microstructures. *Molecules* **2022**, *27*, 2374. [\[CrossRef\]](#)
- Bagheri-Tar, F.; Sahimi, M.; Tsotsis, T.T. Preparation of Polyetherimide Nanoparticles by an Electrospray Technique. *Ind. Eng. Chem. Res.* **2007**, *46*, 3348–3357. [\[CrossRef\]](#)
- Zhang, Z.; Deng, J.; Lu, Y.; Tang, X.; Bao, Y. Dry cutting performance and mechanism of tools with textured composite coatings fabricated by EHDA. *J. Manuf. Process.* **2025**, *156*, 365–380. [\[CrossRef\]](#)
- Wang, J.; Deng, J.; Rong, S.; Zhang, Z.; Bao, Y. Effect of bionic microtexture geometry on the tribological performance of hydrophobic surfaces. *J. Mater. Res. Technol.* **2025**, *38*, 4234–4247. [\[CrossRef\]](#)
- Wang, R.; Deng, J.; Zhang, Z.; Lu, Y.; Li, X.; Ge, D. Preparation and infrared properties of Ni₃Al–Cr₃C₂ composite films deposited by electrohydrodynamic atomization technology. *Mater. Chem. Phys.* **2022**, *278*, 125654. [\[CrossRef\]](#)
- Chen, J.; Wei, M.; Xu, Z.; Wang, Z.; Li, B.; Zhang, W.; Xu, H.; Wang, J.; Pan, J.; Yu, K. Porous Microparticle Preparation via Tuning Electrostatic Breakup Characteristics of a Non-Newtonian Fluid for High-Performance Uranium Extraction. *Ind. Eng. Chem. Res.* **2024**, *63*, 2384–2394. [\[CrossRef\]](#)
- Li, Y.; Wang, Z.; Kong, Q.; Li, B.; Wang, H. Sulfur dioxide absorption by charged droplets in electrohydrodynamic atomization. *Int. Commun. Heat Mass Transf.* **2022**, *137*, 106275. [\[CrossRef\]](#)
- Chen, C.; Liu, W.; Jiang, P.; Hong, T. Coaxial Electrohydrodynamic Atomization for the Production of Drug-Loaded Micro/Nanoparticles. *Micromachines* **2019**, *10*, 125. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yu, D.-G.; Gong, W.; Zhou, J.; Liu, Y.; Zhu, Y.; Lu, X. Engineered shapes using electrohydrodynamic atomization for an improved drug delivery. *WIREs Nanomed. Nanobiotechnol.* **2024**, *16*, e1964. [\[CrossRef\]](#)
- Kim, S.-Y.; Lee, H.; Cho, S.; Park, J.-W.; Park, J.; Hwang, J. Size Control of Chitosan Capsules Containing Insulin for Oral Drug Delivery via a Combined Process of Ionic Gelation with Electrohydrodynamic Atomization. *Ind. Eng. Chem. Res.* **2011**, *50*, 13762–13770. [\[CrossRef\]](#)
- Dau, V.T.; Nguyen, T.-K.; Dao, D.V. Charge reduced nanoparticles by sub-kHz ac electrohydrodynamic atomization toward drug delivery applications. *Appl. Phys. Lett.* **2020**, *116*, 023703. [\[CrossRef\]](#)
- Man, Y.; Zhou, C.; Adhikari, B.; Wang, Y.; Xu, T.; Wang, B. High voltage electrohydrodynamic atomization of bovine lactoferrin and its encapsulation behaviors in sodium alginate. *J. Food Eng.* **2022**, *317*, 110842. [\[CrossRef\]](#)
- Lyu, X.; Wang, X.; Wang, Q.; Ma, X.; Chen, S.; Xiao, J. Encapsulation of sea buckthorn (*Hippophae rhamnoides* L.) leaf extract via an electrohydrodynamic method. *Food Chem.* **2021**, *365*, 130481. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guan, Y.; Sha, Y.; Wu, H.; Zheng, J.; He, B.; Liu, Y.; Lei, Y.; Huang, Y. The spraying characteristics of electrohydrodynamic atomization under different nozzle heights and diameters. *Exp. Therm. Fluid Sci.* **2025**, *169*, 111551. [\[CrossRef\]](#)
- Moreira, K.S.; Di Bonito, L.P.; Glanzer, K.; Carrasco-Munoz, A.; Di Natale, F.; Marques, J.P.M.; Gabriel, P.A.; Oliveira, M.E.; Agostinho, L.L.F. Electric current based automatic classification and operation of EHDA modes. *J. Aerosol Sci.* **2025**, *190*, 106648. [\[CrossRef\]](#)
- Rosell-Llompart, J.; Grifoll, J.; Loscertales, I.G. Electrosprays in the cone-jet mode: From Taylor cone formation to spray development. *J. Aerosol Sci.* **2018**, *125*, 2–31. [\[CrossRef\]](#)
- Wang, J.; Dong, T.; Cheng, Y.; Yan, W.-C. Machine Learning Assisted Spraying Pattern Recognition for Electrohydrodynamic Atomization System. *Ind. Eng. Chem. Res.* **2022**, *61*, 8495–8503. [\[CrossRef\]](#)
- Ma, M.; Zou, Y.; Huang, Z. Deep learning-based automated morphology classification of Electrospun ultrafine fibers from M44 element image of muller matrix. *Optik* **2020**, *206*, 164261. [\[CrossRef\]](#)
- Wang, J.-X.; Wang, X.; Ran, X.; Cheng, Y.; Yan, W.-C. Deep Learning based spraying pattern recognition and prediction for electrohydrodynamic system. *Chem. Eng. Sci.* **2024**, *295*, 120163. [\[CrossRef\]](#)
- Kim, M.J.; Song, J.Y.; Hwang, S.H.; Park, D.Y.; Park, S.M. Electrospray mode discrimination with current signal using deep convolutional neural network and class activation map. *Sci. Rep.* **2022**, *12*, 16281. [\[CrossRef\]](#)
- Sun, J.; Jing, L.; Fan, X.; Gao, X.; Liang, Y.C. Electrohydrodynamic printing process monitoring by microscopic image identification. *Int. J. Bioprint.* **2018**, *5*, 164. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zhou, J.; Shu, X.; Zhang, J.; Yi, F.; Jia, C.; Zhang, C.; Kong, X.; Zhang, J.; Wu, G. A deep learning method based on CNN-BiGRU and attention mechanism for proton exchange membrane fuel cell performance degradation prediction. *Int. J. Hydrogen Energy* **2024**, *94*, 394–405. [\[CrossRef\]](#)
- Jia, C.; He, H.; Zhou, J.; Li, K.; Li, J.; Wei, Z. A performance degradation prediction model for PEMFC based on bi-directional long short-term memory and multi-head self-attention mechanism. *Int. J. Hydrogen Energy* **2024**, *60*, 133–146. [\[CrossRef\]](#)

26. Chai, X.; Zhao, M.; Li, J.; Li, J. Image small target detection in complex traffic scenes based on Yolov8 multiscale feature fusion. *Alex. Eng. J.* **2025**, *126*, 578–590. [[CrossRef](#)]
27. Terven, J.; Córdova-Esparza, D.-M.; Romero-González, J.-A. A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 1680–1716. [[CrossRef](#)]
28. Kang, S.; Hu, Z.; Liu, L.; Zhang, K.; Cao, Z. Object Detection YOLO Algorithms and Their Industrial Applications: Overview and Comparative Analysis. *Electronics* **2025**, *14*, 1104. [[CrossRef](#)]
29. Sirisha, U.; Praveen, S.P.; Srinivasu, P.N.; Barsocchi, P.; Bhoi, A.K. Statistical Analysis of Design Aspects of Various YOLO-Based Deep Learning Models for Object Detection. *Int. J. Comput. Intell. Syst.* **2023**, *16*, 126. [[CrossRef](#)]
30. Vijayakumar, A.; Vairavasundaram, S. YOLO-based Object Detection Models: A Review and its Applications. *Multimed. Tools Appl.* **2024**, *83*, 83535–83574. [[CrossRef](#)]
31. Duan, X.; Wang, P.; Hu, Y.; Li, H.; Yang, S.; Zhu, Y. YOLOv8-DuckPluck: A lightweight target detection model for cherry valley duck feather pecking site detection. *Poult. Sci.* **2025**, *104*, 105484. [[CrossRef](#)] [[PubMed](#)]
32. Wang, Z.; Zhang, Y.; Zhang, S. Real-time personal protective equipment detection and classification with YOLOv8 multi-scale fusion. *J. Real-Time Image Process.* **2025**, *22*, 131. [[CrossRef](#)]
33. Guo, A.; Sun, K.; Zhang, Z. A lightweight YOLOv8 integrating FasterNet for real-time underwater object detection. *J. Real-Time Image Process.* **2024**, *21*, 49. [[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.