

# GED-YOLOv5: An Improved YOLOv5 Model for Weld Defect Detection via Gather-and-Distribute Mechanism and Transfer Learning

Aihua Chen

cah@fzu.edu.cn

Fuzhou University

Tao Ding

Fuzhou University

Peiyuan Zhan

Fuzhou University

Jianfeng Zhong

Fuzhou University

---

## Research Article

**Keywords:** Deep learning, Re-parameterization, Transfer learning, Weld defect detection, YOLOv5

**Posted Date:** May 11th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-8647763/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

**Version of Record:** A version of this preprint was published at Signal, Image and Video Processing on September 1st, 2026. See the published version at <https://doi.org/10.1007/s11760-026-05625-7>.

# GED-YOLOv5: An Improved YOLOv5 Model for Weld Defect Detection via Gather-and-Distribute Mechanism and Transfer Learning

Aihua Chen<sup>1\*</sup>, Tao Ding<sup>1</sup>, Peiyuan Zhan<sup>1</sup>, Jianfeng Zhong<sup>1</sup>

<sup>1</sup>Department of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, 350000, China.

\*Corresponding author(s). E-mail(s): [cah@fzu.edu.cn](mailto:cah@fzu.edu.cn);  
Contributing authors: [TaoD1013@163.com](mailto:TaoD1013@163.com); [zhanpeiyuan2022@163.com](mailto:zhanpeiyuan2022@163.com);  
[zhongjianfeng@fzu.edu.cn](mailto:zhongjianfeng@fzu.edu.cn);

## Abstract

Surface weld defects in circular pipes are often low-contrast, elongated, and easily confused with background texture, which makes accurate real-time inspection challenging on production lines. This work presents GED-YOLOv5, a lightweight enhancement of YOLOv5 tailored for weld-defect localization. Specifically, GED-YOLOv5 integrates (i) a C3\_DSC block with Dynamic Snake Convolution to better model slender and irregular defect patterns, (ii) a re-parameterized Gather-and-Distribute mechanism with efficient fusion to strengthen multi-scale feature aggregation, and (iii) Efficient Channel Attention for low-cost channel recalibration. We further study a transfer-learning pipeline that adapts a model pre-trained on a public TIG weld dataset to a small in-house circular-pipe dataset. Experiments show that GED-YOLOv5 improves F1 and mAP over recent YOLO variants while maintaining real-time throughput (119 FPS on an RTX 4090). Cross-domain evaluation also indicates that source pre-training and fine-tuning are beneficial when target annotations are limited. Overall, GED-YOLOv5 provides a favorable accuracy–efficiency trade-off for industrial weld inspection.

**Keywords:** Deep learning, Re-parameterization, Transfer learning, Weld defect detection, YOLOv5

## 1 Introduction

Welded circular pipes are widely used in construction, petrochemical, marine, and other industrial applications, where weld quality is directly linked to structural integrity and service safety. During manufacturing and in-service operation, surface weld defects such as undercut, cracks, arc craters, and other irregular discontinuities may appear along the weld seam. On real production lines, these defects are often small, low-contrast, and

elongated, and their appearance varies with illumination, surface oxidation, and imaging geometry. These factors make manual inspection subjective and time-consuming and also challenge traditional vision pipelines based on handcrafted features [1–5].

In addition to optical inspection, non-destructive evaluation techniques (e.g., eddy-current testing) have been used to screen metallic surface flaws [6]. Nevertheless, camera-based inspection provides a favorable trade-off among

cost, throughput, and deployment complexity. Deep models have therefore been widely explored for automated surface inspection across domains. Representative studies include steel-surface defect detection [7–10], PCB defect detection under dense components and tiny targets [11–13], and bruise/blemish detection for agricultural produce [14, 15].

Recent works have also demonstrated the effectiveness of YOLO-style detectors for industrial inspection. Zhang *et al.* proposed YOLO-DSCW for seamless stainless-steel pipe surface defect detection [16]. For welding-related vision tasks, Song *et al.* developed an improved YOLO variant to localize weld feature points for robotic welding under strong noise [17]. Attention-enhanced multi-scale frameworks have further been explored to strengthen defect representation under complex backgrounds [18, 19]. These works motivate both accuracy-oriented architectural enhancements and lightweight designs for practical deployment.

With the rapid progress of deep learning, defect inspection has increasingly shifted toward data-driven object detection and segmentation [20, 21]. Meanwhile, weak supervision and anomaly detection have been explored to reduce annotation costs [22–26]. For weld defect inspection, Block *et al.* introduced LoHiWeld to study low-resolution weld inspection under realistic conditions [27], and related works have investigated TIG weld inspection using either classification or detection frameworks [28–30]. Despite promising results, two practical issues remain: (i) publicly available datasets for circular-pipe weld seams are scarce and often cover limited acquisition conditions, and (ii) models trained on one dataset can degrade when applied to new sites due to domain shift [31]. Transfer learning is therefore often adopted to reuse knowledge from related datasets when target annotations are scarce [32, 33].

From a model-design perspective, preserving weak cues and modeling elongated defect patterns are critical for one-stage detectors. Dynamic Snake Convolution (DSC) provides a deformable receptive field that better fits thin or irregular structures [34]. Gold-YOLO proposes an efficient Gather-and-Distribute (GD) mechanism to enhance global-local feature interaction with low computational overhead [35]. In addition, recent IoU variants

(e.g., MPDIoU) can improve localization robustness for objects with varying aspect ratios [36]. These developments motivate us to redesign convolutional and feature aggregation modules for weld defect detection while maintaining a lightweight detector suitable for deployment.

In summary, surface weld defect detection faces the following practical challenges:

1. **Weak and elongated defect patterns:** Small-area and low-contrast defects can be suppressed during feature extraction and multi-scale fusion, leading to missed detections.
2. **Limited and heterogeneous data:** Public datasets for circular-pipe weld inspection are scarce, and data distributions can shift across sites and acquisition setups.
3. **Deployment constraints:** Inspection systems must balance accuracy with real-time throughput and computational cost on production lines [19].

To address these challenges, we propose an improved detector named GED-YOLOv5. The main contributions are summarized as follows:

- We propose GED-YOLOv5, which strengthens multi-scale feature aggregation via an improved GD mechanism with re-parameterized information injection and efficient fusion, thereby preserving weak defect cues.
- We introduce a C3\_DSC module that integrates Dynamic Snake Convolution to better capture slender and irregular defect patterns with modest computational overhead.
- We investigate transfer learning from a public TIG weld dataset (source domain) to a small circular-pipe weld dataset (target domain) and establish a protocol for single-class *bad weld* evaluation.
- Comprehensive experiments, including comparisons with recent YOLO variants, ablation studies, and efficiency analysis, validate the effectiveness of the proposed method.

The remainder of this paper is organized as follows. Section 2 describes the proposed GED-YOLOv5 architecture and the transfer-learning protocol. Section 3 presents experimental settings and results. Section 4 concludes the paper and outlines future work.

## 2 Materials and Methods

YOLOv5 (v6.1) is adopted as the baseline detector due to its favorable accuracy–efficiency trade-off for real-time visual inspection. Nevertheless, surface weld defects on circular pipes are often weak-contrast, slender, and irregular, which can be suppressed during standard convolutional feature extraction and multi-scale fusion. To better accommodate these characteristics, we propose GED-YOLOv5, which augments YOLOv5 with three lightweight components and a transfer-learning protocol for small target-domain datasets: (i) a C3\_DSC block that introduces Dynamic Snake Convolution (DSC) to improve geometric adaptability, (ii) an improved Gather-and-Distribute (GD) feature aggregation mechanism that incorporates re-parameterized information injection and efficient fusion, and (iii) an Efficient Channel Attention (ECA) module to recalibrate channel responses with negligible overhead. An overview of the proposed architecture is provided in Fig. 1. For real-time deployment, lightweight backbone designs such as MobileOne have been investigated to improve latency on edge devices [37].

### 2.1 C3\_DSC Module

Defects such as porosity clusters, undercuts, and lack-of-fusion traces may appear as thin, curved, or fragmented regions. To strengthen the representation of these elongated structures, we replace the standard C3 block with a C3\_DSC module by integrating Dynamic Snake Convolution (DSC) [34]. Unlike rigid grid sampling in standard convolution, DSC learns deformable sampling locations guided by local feature responses, thereby providing a flexible receptive field that better follows defect boundaries.

Let  $X$  denote the input feature map. The DSC-based transformation is expressed as

$$X' = \text{DSCConv}_{1 \times 1}(X), \quad Y' = \text{DSCConv}_{3 \times 3}(X') \quad (1)$$

where  $\text{DSCConv}_{1 \times 1}(\cdot)$  and  $\text{DSCConv}_{3 \times 3}(\cdot)$  denote the DSC operators with  $1 \times 1$  and  $3 \times 3$  kernels, respectively. Following the C3 design, features from multiple bottlenecks are concatenated and projected to form the final output. During training, DSC introduces dynamic kernel modulation; for inference, the effective kernel of each block can be

written as

$$W_{\text{block}} = \alpha \cdot W_{\text{conv}} \odot \text{CDS} + b_{\text{KDS}} \quad (2)$$

$$W_{\text{final}} = \sum_{\text{block}} W_{\text{block}} \quad (3)$$

where  $\alpha$  is an adjustment coefficient,  $b_{\text{KDS}}$  denotes the bias term, and CDS represents the channel-dependent scaling. This formulation allows the module to capture defect geometry with limited additional parameters.

### 2.2 Improved Gather-and-Distribute (GD) Mechanism

Feature aggregation in the neck is critical for detecting defects across scales, especially for small or weak targets that may be suppressed during repeated downsampling. We refine the Gather-and-Distribute (GD) mechanism to improve cross-scale information flow while keeping the model suitable for real-time deployment. The improvements include (a) re-parameterization of the information injection branch to reduce inference latency, (b) a residual path in the fusion module to preserve weak responses, and (c) depthwise separable convolution in pyramid pooling to lower computational overhead.

#### 2.2.1 Re-parameterization of the Information Injection Module

The original information injection module uses multi-branch convolutions to enrich features during training. To avoid the corresponding inference overhead, we fuse the parallel branches into a single equivalent convolution. Specifically, the fused kernel and bias are computed as:

$$K^{\text{fused}} = K^{\text{dense}} + \text{pad}(K^{1 \times 1}) \quad (4)$$

$$b^{\text{fused}} = b^{\text{dense}} + b^{1 \times 1} \quad (5)$$

where  $\text{pad}(\cdot)$  aligns the  $1 \times 1$  kernel to the spatial support of the  $3 \times 3$  kernel. This structural re-parameterization preserves the multi-branch representation during training while deploying a single convolution branch for efficient inference.

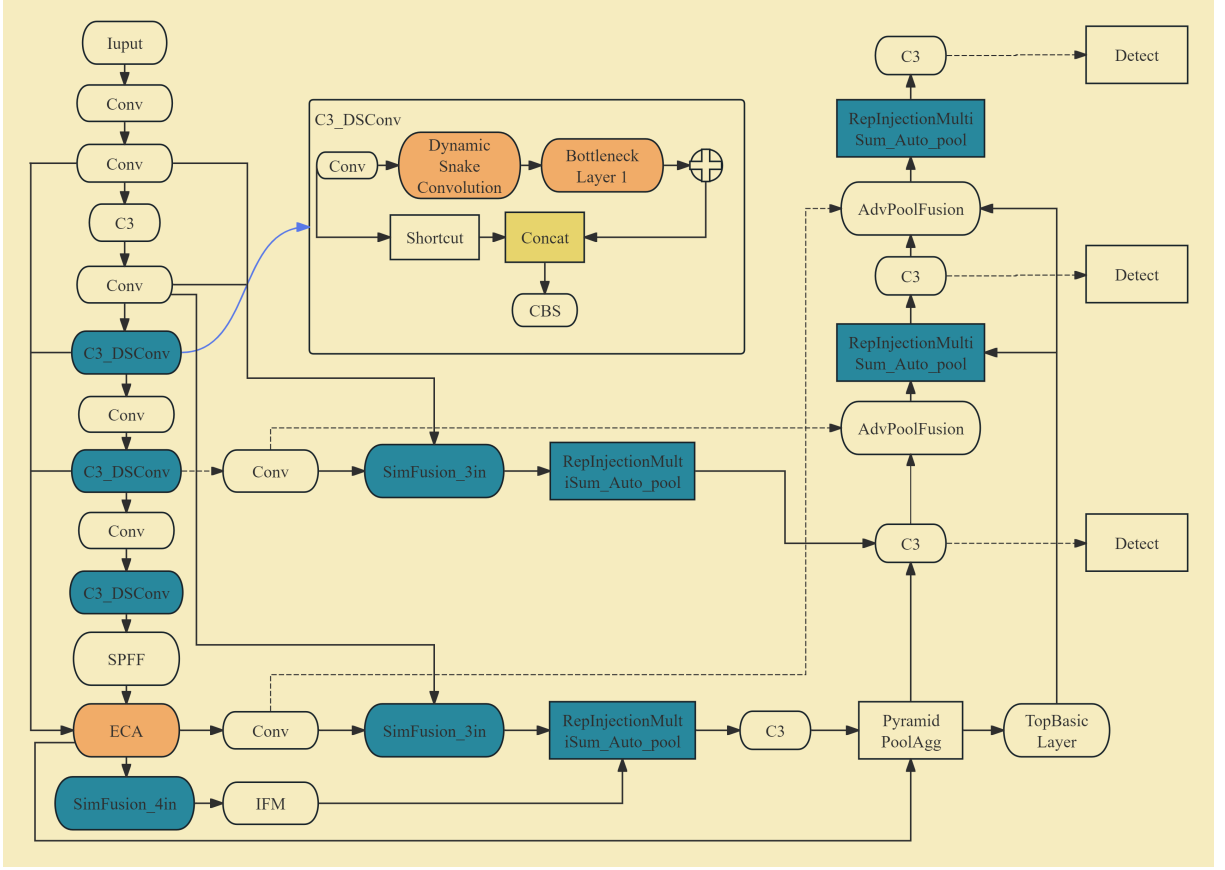


Fig. 1: The overall architecture of the proposed GED-YOLOv5 model.

### 2.2.2 Residual Connection in the Fusion Module

To reduce missed detections on faint defects (e.g., fine porosity), we introduce a residual shortcut that adds a resized intermediate feature map to the fused output:

$$\text{Output} = \text{Fusion}(x_1, x_2, x_3) + \text{Resize}(x_2) \quad (6)$$

This design encourages the fusion module to retain low-level cues that are informative for subtle defect regions.

### 2.2.3 Depthwise Separable Convolution in Pyramid Pooling

Pyramid pooling is effective for enlarging the receptive field but can be computationally expensive when implemented with standard convolutions. We therefore employ depthwise separable convolution in the pyramid pooling stage, reducing parameters

and FLOPs while maintaining multi-scale context modeling.

### 2.3 ECA Attention Mechanism

Weld images often contain strong background texture and illumination variations. To guide the network to emphasize defect-related responses, we incorporate Efficient Channel Attention (ECA) [38] at the end of the backbone. ECA avoids dimensionality reduction and uses a lightweight 1D convolution to capture local cross-channel interactions. The kernel size  $k$  is adaptively determined from the channel dimension  $C$ :

$$k = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (7)$$

where  $\gamma$  and  $b$  are hyper-parameters and  $(\cdot)_{\text{odd}}$  denotes the nearest odd integer. The resulting channel weights are applied to recalibrate the

feature map with minimal computational cost. Additional ECA design details are provided in the Supplementary Material.

## 2.4 Transfer Learning for Small Sample Datasets

In practical deployments, collecting and labeling large-scale circular-pipe weld defect datasets can be costly. We therefore adopt a transfer-learning strategy that leverages a public TIG welding dataset (source domain) for pre-training [28]. The model is first trained on the source dataset with multi-class defect annotations to learn generic weld appearance features. Then, the learned weights are used to initialize training on the in-house circular-pipe dataset (target domain), where only limited samples are available. This procedure improves optimization stability and reduces overfitting compared with training from scratch. Representative source/target examples are provided in the Supplementary Material.

# 3 Experiments and Discussion

## 3.1 Experimental Setup

All experiments were performed on an Ubuntu workstation equipped with an NVIDIA GeForce RTX 4090 GPU (24 GB). Table 1 summarizes the experimental configuration and the default training settings. Unless otherwise specified, all models were trained and evaluated under the same protocol to ensure a fair comparison.

### *Inference protocol and speed measurement.*

All quantitative results in Table 3 are reported on the same *test* split with batch size = 1 and identical input resolution. Unless otherwise noted, we follow the official YOLOv5 validation script (val.py) with `conf_thres=0.001` and NMS IoU threshold 0.6 for metric computation. Throughput (FPS) is measured using the official evaluation pipeline, including preprocessing and non-maximum suppression (NMS) but excluding disk I/O. We discard the first 50 warm-up iterations and report the average over the full test set. Half-precision (FP16) inference is enabled when supported by the backend.

### *Preprocessing and augmentation.*

Images were normalized using channel-wise mean and variance. We applied standard YOLOv5 data augmentation, including mosaic augmentation and random horizontal flips. Generative data augmentation (e.g., GAN-based synthesis) has also been explored to alleviate data scarcity in visual inspection [25, 39]. To improve robustness on the small target-domain dataset, mild geometric perturbations (rotation within  $\pm 10^\circ$ ) and random cropping were additionally used.

### 3.1.1 Dataset and Preprocessing

We evaluate GED-YOLOv5 on two domains: a *source-domain* public TIG welding dataset and a *target-domain* in-house circular-pipe weld defect dataset. The source dataset follows the VOC format and provides multi-class annotations of weld appearance/defect regions. The target dataset contains images captured by an industrial camera under controlled illumination; each image includes one or more defect regions annotated with bounding boxes. Details of the acquisition setup are provided in the Supplementary Material. For discontinuous defects, multiple boxes are used to cover separated defect segments. Dataset splits are summarized in Table 2. The target-domain dataset contains 170 *raw* images. We randomly split the raw target images into 120/25/25 for train/val/test. To mitigate data scarcity, we perform *offline* augmentation on the training split, expanding 120 raw training images to 600 images, while keeping the validation and test sets unchanged (25 images each) for an unbiased evaluation. The resulting target splits contain 790, 28, and 35 defect bounding boxes for train(aug)/val/test, respectively.

### *Experimental settings.*

To connect the public multi-class benchmark with the practical in-house inspection scenario, we report two experimental settings:

- **Source-domain in-domain evaluation.** Train, validate, and test on the public TIG weld defect dataset with the full label set (six defect types plus defect-free samples). This setting supports a controlled comparison among different YOLO variants (Table 3).
- **Source→target zero-shot evaluation.** Train on the public dataset and directly evaluate

**Table 1:** Experimental configuration and default training settings.

| Item             | Setting                                                |
|------------------|--------------------------------------------------------|
| Operating system | Ubuntu                                                 |
| GPU              | NVIDIA GeForce RTX 4090 (24 GB)                        |
| Input size       | $640 \times 640$                                       |
| Batch size       | 16 (source domain); 8 (target domain)                  |
| Framework        | PyTorch 1.11.0                                         |
| Optimizer        | SGD (momentum 0.937, weight decay $5 \times 10^{-4}$ ) |
| Schedule         | 200 epochs, warm-up 3 epochs + cosine annealing        |

**Table 2:** Dataset composition and splits (number of images). For the target domain, the training split contains 120 raw images and is expanded to 600 images after offline augmentation (shown in parentheses); validation and test sets remain unchanged. Target-domain boxes: 790/28/35 for train(aug)/val/test.

| Dataset             | Train     | Val | Test |
|---------------------|-----------|-----|------|
| Source domain total | 890       | 255 | 127  |
| Target domain total | 120 (600) | 25  | 25   |

on the target-domain test split without any parameter updates. Because the target-domain annotations focus on pass/fail inspection, we collapse all predicted defect categories into a single *bad weld* class for evaluation (Table 5).

### 3.1.2 Training Details

We trained models using PyTorch 1.11.0 and the official YOLOv5 training pipeline. Stochastic Gradient Descent (SGD) was adopted with momentum 0.937 and weight decay  $5 \times 10^{-4}$ . The initial learning rate was 0.01 with a linear warm-up during the first three epochs, followed by cosine annealing. Unless otherwise stated, training was conducted for 200 epochs.

For cross-domain evaluation in the source→target setting, the detector is trained on the source-domain TIG dataset and then directly tested on the target-domain split without further adaptation. Therefore, no additional fine-tuning schedules (e.g., freezing/unfreezing) are used unless explicitly stated.

### 3.1.3 Evaluation Metrics

We report detection effectiveness using precision ( $P$ ), recall ( $R$ ), F1-score, and mean Average Precision (mAP). Let  $TP$ ,  $FP$ , and  $FN$  denote the numbers of true positives, false positives, and false negatives, respectively. Precision and recall are defined as

$$P = \frac{TP}{TP + FP}, \quad R = \frac{TP}{TP + FN}. \quad (8)$$

The F1-score is the harmonic mean of  $P$  and  $R$ :

$$F1 = \frac{2PR}{P + R}. \quad (9)$$

Following common practice in object detection, we compute mAP@0.5 at an IoU threshold of 0.5, and COCO-style mAP@0.5:0.95 by averaging over IoU thresholds from 0.50 to 0.95 with a step of 0.05. Efficiency is reported as end-to-end latency and frames per second (FPS) measured on the same hardware.

## 3.2 Results and Discussion

### 3.2.1 Overall Performance of GED-YOLOv5

We first analyze GED-YOLOv5 on the source-domain TIG dataset. Compared with the YOLOv5 baseline, the proposed model exhibits a consistently better precision–recall trade-off and improved robustness to background clutter and weak-contrast defect patterns. The improvements are more pronounced on categories with slender or fragmented morphology, aligning with the design motivation of the refined feature aggregation and the C3\_DSC replacement. In addition, qualitative inspections indicate that GED-YOLOv5 reduces missed detections on small/low-contrast defects

and suppresses spurious predictions under complex weld textures. For completeness, we provide PR curves, per-class breakdowns, and additional qualitative examples in the Supplementary Material.

### 3.2.2 Comparison with State-of-the-Art Models

Table 3 compares GED-YOLOv5 with several recent YOLO variants in terms of model size, accuracy, and speed. We compare against YOLOv7 and the small variants YOLOv8s/YOLOv9s/YOLOv10s using publicly available implementations (official repositories where available). Unless stated otherwise, all baselines are trained from scratch on the same data split with identical input resolution (640) and the same training schedule (200 epochs), following the default training recipes of each framework. GED-YOLOv5 achieves the highest mAP@0.5 and F1-score among the compared models while keeping the parameter count in a similar range to lightweight detectors. Importantly, the inference speed remains competitive, which is essential for in-line inspection.

#### *Discussion of quantitative comparison.*

Table 3 compares GED-YOLOv5 with representative YOLO variants under the unified protocol described in Section 3.1. GED-YOLOv5 achieves the best F1-score (91%) and the highest mAP@0.5 (91.3%) while keeping the parameter count low (9.2M) and maintaining real-time throughput (119FPS). Under the stricter COCO-style metric mAP@0.5:0.95, YOLOv8s reaches the highest value (75.3%), indicating stronger localization performance at high IoU thresholds, whereas GED-YOLOv5 attains 70.3% with a substantially higher inference speed. Compared with YOLOv7, which yields a competitive mAP@0.5:0.95 (74.1%) but requires 37.2M parameters, GED-YOLOv5 offers a more favorable accuracy–efficiency trade-off for in-line inspection. Overall, the results suggest that the proposed architectural modifications primarily improve detection reliability (precision/recall balance) and coarse-to-medium localization accuracy, while further gains under high-IoU evaluation may benefit from higher-resolution training and boundary regression, which we consider in future work.

### 3.2.3 Training Curves and Convergence

GED-YOLOv5 exhibits stable optimization behavior under the same training schedule as the baseline, without introducing noticeable oscillations. Training-loss trajectories and mAP evolution curves are provided in the Supplementary Material for reference.

### 3.3 Ablation Study

We conduct an ablation study on the public TIG weld dataset to quantify the contribution of each proposed component under identical training and evaluation settings. To keep the presentation concise for the two-column layout, Table 4 reports the impact of enabling (i) the optimized Global-Dependency (GOLD) module (GD), (ii) the C3\_DSC replacement, and (iii) the ECA attention. We summarize performance using F1 and mAP metrics (mAP@0.5 and the stricter COCO-style mAP@0.5:0.95). Here, F1 is computed from precision and recall reported by the evaluation script as  $F1 = \frac{2PR}{P+R}$ . Starting from the YOLOv5s baseline, each individual component yields a consistent gain in mAP@0.5. The full combination (GED-YOLOv5) achieves the best overall accuracy on this benchmark, while detailed model size comparisons for the final models are provided in Table 3.

### 3.4 Transfer Learning Results

Table 5 reports target-domain single-class results (bad weld) under three protocols: (A) *zero-shot* transfer (directly evaluating the source-trained detector on the target set without any target fine-tuning), (B) *source-initialized fine-tuning* on the target training set, and (C) *target-only* training from random initialization.

All protocols use the same target-domain split in Table 2. For protocol (B), we initialize the model with the corresponding source-domain checkpoint and fine-tune all layers (freeze=0) for 200 epochs using SGD with  $lr_0 = 0.01$  ( $lrf = 0.01$ ), momentum 0.937, and weight decay  $5 \times 10^{-4}$ , with input size  $640 \times 640$  and batch size 8. Since the target task is single-class, the detection head is adapted to  $nc=1$ ; when  $nc$  changes, the YOLOv5 training pipeline transfers backbone/neck weights and re-initializes the detection head.

**Table 3:** Performance comparison on the public TIG weld dataset (source domain). All models are trained and evaluated with input size  $640 \times 640$ ; FPS is measured on an RTX 4090 with FP16 inference and batch size 1.

| Model             | Params (M) | F1 (%)      | mAP@0.5 (%) | mAP@0.5:0.95 (%) | FPS          |
|-------------------|------------|-------------|-------------|------------------|--------------|
| YOLOv5            | 7.0        | 88.0        | 88.6        | 68.9             | 120.5        |
| YOLOv7            | 37.2       | 90.0        | 90.5        | 74.1             | 116.1        |
| YOLOv8s           | 11.1       | 89.0        | 91.0        | 75.3             | 86.2         |
| YOLOv9s           | 9.7        | 81.0        | 84.8        | 62.7             | 80.0         |
| YOLOv10s          | 8.0        | 83.0        | 86.0        | 68.3             | 147.0        |
| <b>GED-YOLOv5</b> | <b>9.2</b> | <b>91.0</b> | <b>91.3</b> | <b>70.3</b>      | <b>119.0</b> |

**Table 4:** Ablation study of the proposed modules on the public TIG weld dataset. All variants are trained with identical settings.

| Setting            | GD | C3_DSC | ECA | F1          | mAP@0.5     | mAP@0.5:0.95 |
|--------------------|----|--------|-----|-------------|-------------|--------------|
| Baseline (YOLOv5s) | -  | -      | -   | 88.0        | 88.6        | 68.9         |
| +GD (GOLD)         | ✓  | -      | -   | 89.1        | 89.8        | 69.1         |
| +ECA               | -  | -      | ✓   | 88.6        | 89.2        | 68.9         |
| +C3_DSC            | -  | ✓      | -   | 87.7        | 89.8        | 65.0         |
| +ECA+C3_DSC        | -  | ✓      | ✓   | 89.4        | 89.6        | 69.6         |
| Full (GED-YOLOv5)  | ✓  | ✓      | ✓   | <b>91.0</b> | <b>91.3</b> | <b>70.3</b>  |

In the zero-shot setting (A), **GED-YOLOv5** improves recall from 22.9% to 45.7% and F1 from 28.3% to 37.7%, indicating better cross-domain generalization. With source-initialized fine-tuning (B), **GED-YOLOv5** achieves 63.8% precision / 51.4% recall / 56.9% F1 and 51.7% mAP@0.5, outperforming YOLOv5 (59.6% / 40.0% / 47.9% and 43.9% mAP@0.5), while remaining comparable

on mAP@0.5:0.95 (20.3% vs. 21.1%). For target-only training (C), **GED-YOLOv5** yields a large margin (P/R/F1 = 80.4/82.2/81.3%, mAP@0.5 = 86.2%, mAP@0.5:0.95 = 55.9%) compared with YOLOv5 (61.7/37.1/46.3%, 43.3%, 14.5%). Overall, GED-YOLOv5 consistently improves target-domain detection robustness under domain shift and limited target data.

**Table 5:** Target-domain single-class evaluation under protocols (A)–(C). The target dataset uses *bad weld* as the only defect class.

| Model      | Protocol / initialization              | Precision (%) | Recall (%)  | F1 (%)      | mAP@0.5 (%) | mAP@0.5:0.95 (%) |
|------------|----------------------------------------|---------------|-------------|-------------|-------------|------------------|
| YOLOv5     | (A) source→target, zero-shot           | 37.2          | 22.9        | 28.3        | 27.0        | 9.7              |
| GED-YOLOv5 | (A) source→target, zero-shot           | 32.1          | 45.7        | 37.7        | 25.7        | 6.4              |
| YOLOv5     | (B) fine-tune on target (source init.) | 59.6          | 40.0        | 47.9        | 43.9        | <b>21.1</b>      |
| GED-YOLOv5 | (B) fine-tune on target (source init.) | <b>63.8</b>   | <b>51.4</b> | <b>56.9</b> | <b>51.7</b> | 20.3             |
| YOLOv5     | (C) target-only training               | 61.7          | 37.1        | 46.3        | 43.3        | 14.5             |
| GED-YOLOv5 | (C) target-only training               | <b>80.4</b>   | <b>82.2</b> | <b>81.3</b> | <b>86.2</b> | <b>55.9</b>      |

### 3.5 Runtime and Reproducibility

GED-YOLOv5 runs at 119.0 FPS on the RTX 4090 GPU under the evaluation settings described above, which is comparable to the baseline YOLOv5 (Table 3). To support reproducibility, we will provide training scripts, hyper-parameters, and pre-trained weights upon reasonable request, together with the finalized dataset split configuration.

## 4 Conclusion

This paper presented GED-YOLOv5, a lightweight detector for surface weld defect localization on circular pipes. By integrating the proposed C3\_DSC block with Dynamic Snake Convolution, an improved Gather-and-Distribute feature aggregation mechanism, and an Efficient Channel Attention module, the detector better preserves weak cues and captures slender defect patterns with limited overhead. On the public TIG weld dataset, GED-YOLOv5 achieves 91.0% F1, 91.3% mAP@0.5, and 70.3% mAP@0.5:0.95 while running at 119 FPS with 9.2 M parameters (Table 3). Cross-domain evaluation on a small in-house circular-pipe dataset further verifies the generalization performance of the source-trained models under domain shift (Table 5).

Future work will focus on (1) improving robustness to high-illumination reflections and severe background texture, (2) expanding the in-house dataset with more diverse operating conditions and defect types, and (3) investigating domain generalization strategies to mitigate performance degradation caused by acquisition and sensor variations in practical production lines.

## Declarations

- Funding: –
- Competing interests: The authors declare that they have no competing interests.
- Ethics approval: Not applicable.
- Consent to participate: Not applicable.
- Consent for publication: Not applicable.
- Data availability: The source-domain TIG weld defect dataset used in this study is publicly available online (see the dataset citation in the References). The in-house circular-pipe dataset contains proprietary industrial data and is

available from the corresponding author upon reasonable request.

- Code availability: The implementation and trained models will be made available at – (or from the corresponding author upon reasonable request).
- Materials availability: Not applicable.
- Author contributions: Conceptualization: Aihua Chen; Methodology: Tao Ding; Data curation: Tao Ding, Peiyuan Zhan; Writing–original draft: Tao Ding; Writing–review & editing: Aihua Chen, Jianfeng Zhong; Supervision: Aihua Chen.

## References

- [1] Zhang, S.: Industry 4.0 and intelligent manufacturing. *Mech. Des. Manuf. Eng.* 43(8), 1–5 (2014)
- [2] Hussen, M.S., Kyosev, Y.K., Pietsch, K., Boll, J., Kabish, A.K.: Estimation of the size and location of an artificial weld seam defect based on the actual ultrasonic welding phenomenon for online inspection. *Physica Scripta* 98(8), 085921 (2023) <https://doi.org/10.1088/1402-4896/ace390>
- [3] Mao, H., Ren, J., Tang, Y., Mao, H., Chen, Y., Yi, X., Huang, Z., Li, X.: Detection of weld defects using ultrasonic-guided waves based on matching pursuit and density peak clustering. *Transactions of the Institute of Measurement and Control* 45(8), 1470–1483 (2023) <https://doi.org/10.1177/01423312221140618>
- [4] Ajmi, C., Zapata, J., Elferchichi, S., Laabidi, K.: Advanced faster-rcnn model for automated recognition and detection of weld defects on limited x-ray image dataset. *Journal of Nondestructive Evaluation* 43(1), 14 (2024) <https://doi.org/10.1007/s10921-023-01032-x>
- [5] He, Z., Pei, Z., Li, E., Zhou, E., Huang, Z., Xing, Z., Li, B.: An image segmentation based localization method for detecting weld seams. *Advances in Engineering Software* 194, 103662 (2024) <https://doi.org/10.1016/j.advengsoft.2024.103662>
- [6] Wang, C., Wang, K., Cong, Z., Cui, Z.: Eddy current testing on weld defect based on the dual frequency independent component analysis. In: 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), pp. 1–5 (2017). <https://doi.org/10.1109/I2MTC.2017.8288888>

- org/10.1109/I2MTC.2017.7969919
- [7] Wen, X., Shan, J., He, Y., Song, K.: Steel surface defect recognition: A survey. *Coatings* 13(1), 17 (2022) <https://doi.org/10.3390/coatings13010017>
  - [8] Zhao, C., Shu, X., Yan, X., Zuo, X., Zhu, F.: RDD-YOLO: A modified yolo for detection of steel surface defects. *Measurement* 214, 112776 (2023) <https://doi.org/10.1016/j.measurement.2023.112776>
  - [9] Liu, R., Huang, M., Gao, Z., Cao, Z., Cao, P.: MSC-DNet: An efficient detector with multi-scale context for defect detection on strip steel surface. *Measurement* 209, 112467 (2023) <https://doi.org/10.1016/j.measurement.2023.112467>
  - [10] Mao, W.-L., Wang, C.-C., Chou, P.-H., Liu, Y.-T.: Automated defect detection for massproduced electronic components based on yolo object detection models. *IEEE Sensors Journal* 24(16), 26877–26888 (2024) <https://doi.org/10.1109/JSEN.2024.3418618>
  - [11] Lv, S., Ouyang, B., Deng, Z., Liang, T., Jiang, S., Zhang, K., Chen, J., Li, Z.: A dataset for deep learning based detection of 15 printed circuit board surface defect. *Scientific Data* 11(1), 811 (2024) <https://doi.org/10.1038/s41597-024-03656-8>
  - [12] Chen, X., Wu, Y., He, X., Ming, W.: A comprehensive review of deep learning-based pcb defect detection. *IEEE Access* 11, 139017–139038 (2023) <https://doi.org/10.1109/ACCESS.2023.3339561>
  - [13] Chen, C., Wu, Q., Zhang, J., Xia, H., Lin, P., Wang, Y., Tian, M., Song, R.: U 2 D 2 PCB: Uncertainty-aware unsupervised defect detection on pcb images using reconstructive and discriminative models. *IEEE Transactions on Instrumentation and Measurement* 73, 1–10 (2024) <https://doi.org/10.1109/TIM.2024.3386210>
  - [14] Fan, S., Liang, X., Huang, W., Zhang, V.J., Pang, Q., He, X., Li, L., Zhang, C.: Real-time defects detection for apple sorting using nir cameras with pruning-based yolov4 network. *Computers and Electronics in Agriculture* 14, 193, 106715 (2022) <https://doi.org/10.1016/j.compag.2022.106715>
  - [15] Unal, Z., Kızıldeniz, T., Özden, M., Aktaş, H., Karagöz, O.: Detection of bruises on red apples using deep learning models. *Scientia Horticulturae* 329, 113021 (2024) <https://doi.org/10.1016/j.scienta.2024.113021>
  - [16] Zhang, C., Zhang, J., Wang, Y., Wang, H., Wang, Y.: YOLO-DSCW: a seamless stainless steel pipe surface defect detection model. *Signal, Image and Video Processing* 19, 1180 (2025) <https://doi.org/10.1007/s11760-025-04805-1>
  - [17] Song, L., Kang, J., Zhang, Q., Wang, S.: A weld feature points detection method based on improved YOLO for welding robots in strong noise environment. *Signal, Image and Video Processing* 17, 1801–1809 (2023) <https://doi.org/10.1007/s11760-022-02391-0>
  - [18] Xu, J., Hu, X., Zhang, Y.: TGA-YOLO: An enhanced detector for multi-scale industrial defect detection with dual attention mechanisms. *Signal, Image and Video Processing* 19, 931 (2025) <https://doi.org/10.1007/s11760-025-04526-5>
  - [19] Zhou, M., Wang, R., Chen, Y.: PMSE-YOLO: an efficient framework for detecting surface defects of hot-rolled strip steel. *Signal, Image and Video Processing* 19, 534 (2025) <https://doi.org/10.1007/s11760-025-04189-2>
  - [20] Qian, W., Dong, S., Chen, L., Ren, Q.: Image enhancement method for low-light pipeline weld x-ray radiographs based on weakly supervised deep learning. *NDT & E International* 143, 103049 (2024) <https://doi.org/10.1016/j.ndteint.2024.103049>
  - [21] Huang, L., Liao, C., Song, X., Chen, T., Zhang, X., Deng, Z.: Research on detection mechanism of weld defects of carbon steel plate based on orthogonal axial eddy current probe. *Sensors* 20(19), 5515 (2020) <https://doi.org/10.3390/s20195515>
  - [22] Tabernik, D., Šela, S., Skvarč, J., Skočaj, D.: Segmentation-based deeplearning approach for surface-defect detection. *Journal of Intelligent Manufacturing* 31(3), 759–776 (2020) <https://doi.org/10.1007/s10845-019-01476-x>
  - [23] Tsai, D.-M., Fan, S.-K.S., Chou, Y.-H.: Auto-annotated deep segmentation for surface defect detection. *IEEE Transactions on Instrumentation and Measurement* 70, 1–10 (2021) <https://doi.org/10.1109/TIM.2021.3087826>
  - [24] Huang, F., Wang, B.-w., Li, Q.-p., Zou, J.: Texture surface defect detection of plastic relays with an enhanced feature pyramid network. *Journal of Intelligent Manufacturing*

- 34(3), 1409–1425 (2023) <https://doi.org/10.1007/s10845-021-01864-2>
- [25] Zhang, H., Pan, D., Liu, J., Jiang, Z.: A novel mas-gan-based data synthesis method for object surface defect detection. *Neurocomputing*499, 106–114 (2022) <https://doi.org/10.1016/j.neucom.2022.05.021>
- [26] Zhang, Z., Shang, Z., Wang, X., Ma, J.: Combined anomaly aware weakly supervised lightweight model for surface defect inspection. *IEEE Transactions on Industrial Informatics*20(4), 6652–6663 (2024) <https://doi.org/10.1109/TII.2023.3348835>
- [27] Block, S.B., Silva, R.D., Lazzaretti, A.E., Minetto, R.: LoHi-WELD: A novel industrial dataset for weld defect detection and classification, a deep learning study, and future perspectives. *IEEE Access*12, 77442–77453 (2024) <https://doi.org/10.1109/ACCESS.2024.3407019>
- [28] Bacioiu, D., Melton, G., Papaelias, M., Shaw, R.: Automated defect classification of aluminium 5083 tig welding using hdr camera and neural networks. *Journal of manufacturing processes*45, 603–613 (2019) <https://doi.org/10.1016/j.jmapro.2019.07.020>
- [29] Sonwane, S., Chiddarwar, S.: Developing adss for enhancing weldment defect detection, classification, and remediation using hdr images and adaptive mdcnnet neural network. *Journal of Nondestructive Evaluation*43(1), 16 (2024) <https://doi.org/10.1007/s10921-023-01027-8>
- [30] Xu, L., Dong, S., Wei, H., Peng, D., Qian, W., Ren, Q., Wang, L., Ma, Y.: Intelligent identification of girth welds defects in pipelines using neural networks with attention modules. *Engineering Applications of Artificial Intelligence*127, 107295 (2024) <https://doi.org/10.1016/j.engappai.2023.107295>
- [31] Wang, J., Zhang, Q., Ding, C., Ren, Y., Chu, J., Wang, H., Zhu, Y.: Detection and evaluation of dissimilar metal weld defects based on the tx-rx pulsed eddy current testing probe. *Russian Journal of Nondestructive Testing*60(3), 306–317 (2024) <https://doi.org/10.1134/S1061830924600096>
- [32] Kumaresan, S., Aultrin, K.S.J., Kumar, S.S., Anand, M.D.: Transfer learning with cnn for classification of weld defect. *IEEE Access*9, 95097–95108 (2021) <https://doi.org/10.1109/ACCESS.2021.309348716>
- [33] Kumaresan, S., Aultrin, K.J., Kumar, S., Anand, M.D.: Deep learning-based weld defect classification using vgg16 transfer learning adaptive fine-tuning. *International Journal on Interactive Design and Manufacturing (IJI-DeM)*17(6), 2999–3010 (2023) <https://doi.org/10.1007/s12008-023-01327-3>
- [34] Qi, Y., He, Y., Qi, X., Zhang, Y., Yang, G.: Dynamic snake convolution based on topological geometric constraints for tubular structure segmentation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6070–6079 (2023). <https://doi.org/10.1109/ICCV51070>. 2023.00559
- [35] Wang, C., He, W., Nie, Y., Guo, J., Liu, C., Wang, Y., Han, K.: Gold-YOLO: Efficient object detector via gather-and-distribute mechanism. *Advances in Neural Information Processing Systems*36, 51094–51112 (2023) <https://doi.org/10.48550/arXiv.2309.11331>
- [36] Ma, S., Xu, Y.: MPDIoU: A loss for efficient and accurate bounding box regression. *arXiv preprint* (2023) <https://doi.org/10.48550/arXiv.2307.07662> arXiv:2307.07662
- [37] Vasu, P.K.A., Gabriel, J., Zhu, J., Tuzel, O., Ranjan, A.: Mobileone: An improved one millisecond mobile backbone. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7907–7917 (2023). <https://doi.org/10.48550/arXiv.2206.04040>
- [38] Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q.: ECA-Net: Efficient channel attention for deep convolutional neural networks. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11531–11539 (2020). <https://doi.org/10.1109/CVPR42600.2020.01155>
- [39] Guo, R., Liu, H., Xie, G., Zhang, Y.: Weld defect detection from imbalanced radiographic images based on contrast enhancement conditional generative adversarial network and transfer learning. *IEEE Sensors Journal*21(9), 10844–10853 (2021) <https://doi.org/10.1109/JSEN.2021.3059860>

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementary.zip](#)