Bio-FlameNet: Flame Detection for Biogas Digesters

Xinzhe Yue[1], Runqian Zhang[1], Yingrui Geng[1], Mengtao Wang[1],
Zenghui Wang[2], Song Wang[3], Lin Meng[1,*]
{gr0794hp@ed.ritsumei.ac.jp, gr0663sx@ed.ritsumei.ac.jp, gr0731rp@ed.ritsumei.ac.jp}
[1] *Graduate School of Science and Engineering, Ritsumeikan University, 525-8577, Shiga, Japan*
[2] *Department of Electrical Engineering, University of South Africa, Florida 1710, South Africa*
[3] *Kansai Gaidai College, 573-1001, Osaka, Japan*
[*] *Corresponding author*

Abstract. Biogas facilities are crucial for converting organic waste into renewable energy, yet their inherent methane-related fire and explosion risks pose significant operational safety threats. Deploying automated early-flame detection systems in this high-risk industrial setting faces two core challenges. First, there is extreme scarcity of domain-specific real-world fire data due to safety constraints. Second, the detector must alarm at the incipient stage of a fire, where the target is extremely small. This paper proposes Bio-FlameNet, a data-centric framework for building an efficient and reliable flame detector in the absence of real-world fire data. We generate a high-fidelity synthetic dataset by fusing flame assets onto real biogas-facility backgrounds using a copy–paste pipeline, explicitly controlling the small:medium:large flame ratio to 7:2:1. We further adopt a two-stage transfer-learning strategy by pre-training on D-Fire and fine-tuning on domain-specific synthetic data. Experiments on the YOLOv8 family show that the proposed approach improves lightweight-model performance while substantially reducing training time across all model scales.

Keywords: Biogas Safety; Flame Detection; Small Object Detection; Synthetic Data; Transfer Learning.

1 Introduction

Anaerobic Digestion (AD) is a cornerstone technology for renewable energy production, efficiently converting organic waste into methane-rich biogas. However, this process inherently creates a high-risk operational environment. Biogas typically contains 40% to 70% methane (CH_4), which has a Lower Explosive Limit (LEL) of just 5%[1]. This means that if a leak occurs from the digester, pipes, or valves and encounters a potential ignition source (such as electrical equipment, static sparks, or illicit smoking), it can easily lead to a fire or even a catastrophic explosion[2]. Therefore, implementing 24/7, automated, real-time fire monitoring for biogas facilities—capable

IIKI 2025, December 18-21, Chengdu, People's Republic of China

DOI 10.4108/eai.18-12-2025.2365245

of issuing an alert at the very earliest stage of a fire (i.e., when a small flame appears)—is a core requirement for protecting personnel and assets.

Traditional fire alert systems rely on point sensors, such as smoke, heat, or gas detectors[3]. However, for large-scale, open, or semi-open industrial environments like biogas digesters, these sensors have inherent limitations. They are passive, requiring smoke, heat, or flammable gas to diffuse to the sensor's location to trigger an alarm. This creates spatial coverage blind spots and significant time delays in response.

Computer vision-based monitoring systems offer a highly promising alternative. A vision system provides wide-area, non-contact, real-time monitoring via cameras, enabling detection the instant a flame appears (i.e., when the light signal is captured).

However, directly applying general-purpose deep learning object detection models (like YOLO or Faster R-CNN) to a biogas digester scene is unfeasible. These models are data-driven, and public flame datasets (which mostly consist of forest or urban building fires) have scene characteristics that are drastically different from the industrial environment of a biogas digester. Models trained on such data suffer from poor generalization, resulting in high false-positive and false-negative rates in real-world applications.

Deploying an intelligent vision system in a biogas environment requires navigating a complex set of interconnected obstacles. The primary hurdle is the extreme scarcity of data; because biogas digester fires are low-probability yet high-risk events, publicly available annotated data for this specific setting is virtually non-existent, and safety constraints make it infeasible to replicate real fires for data collection. This challenge is compounded by high scene specificity, as digesters are cluttered with reflective metal pipes and valves that produce strong highlights under varying lighting, easily mimicking the visual features of early-stage flames and leading to severe false alarms. Ultimately, the system must satisfy an urgent early detection requirement, identifying the initial flame[4] when it occupies only a tiny fraction of the image. Technologically, this constitutes a highly demanding Small Object Detection (SOD) problem[5], where critical visual information is prone to being lost during deep feature propagation. These compounded difficulties render off-the-shelf detection models ineffective, creating a critical need for a specialized framework designed to function in the absence of real-world training data.

To address all of these challenges, this paper proposes Bio-FlameNet, a complete, data-centric framework to systematically build a flame detection model specialized for the biogas digester environment.

The core contributions of this work are:

- **Data Level:** To address data scarcity and scene specificity, this paper proposes a high-fidelity synthetic data generation pipeline that fuses flame assets with real biogas digester background images.
- **Algorithmic Level:** To tackle the early detection (SOD) problem, this paper introduces mathematical size distribution control during the data generation phase. This method allows for precise control over the size distribution of flames in the synthetic dataset, enforcing a 7:2:1 ratio of small:medium:large flames to "force" the model to focus on learning small target detection.

- **Training Strategy:** This paper adopts a two-stage knowledge transfer strategy. First, the model is pre-trained on the large D-Fire dataset[6] to learn the general features of fire. Then, it is fine-tuned on the domain-specific synthetic dataset to adapt to the special characteristics of the biogas scene.
- **Empirical Evaluation:** Using the full series of YOLOv8 models, this paper provides a comprehensive empirical evaluation and in-depth performance analysis of the proposed transfer learning strategy against a "train from scratch" baseline, offering key data on the efficiency and accuracy trade-offs for practical deployment.

The remainder of this paper is organized as follows: Section 2 reviews related work. Section 3 details the methodology of Bio-FlameNet. Section 4 provides the experimental setup, quantitative, and qualitative evaluations. Section 5 concludes the paper and discusses future work.

2 RELATED WORK

2.1 Vision-Based Flame and Smoke Detection

Early vision-based fire detection relied on hand-crafted features. Researchers analyzed the physical properties of flames and smoke, utilizing their unique color features in RGB or HSI/HSV color spaces[7], combined with dynamic features like motion analysis, flicker frequency, and edge irregularity[7]. While computationally inexpensive, these methods lacked robustness and suffered from extremely high false alarm rates in complex lighting or in the presence of distractors (e.g., red objects, lights).

The rise of deep learning, particularly the success of Convolutional Neural Networks (CNNs)[8], fundamentally changed this field. Modern approaches treat fire detection as an object detection problem, using state-of-the-art frameworks like Faster R-CNN or the YOLO (You Only Look Once) series. These models can automatically learn deep, abstract features of flames and smoke from data, achieving detection accuracy far surpassing traditional methods.

2.2 Synthetic Data and Data Augmentation

The "data-hungry" nature of deep learning models is a primary bottleneck for their application in many data-scarce, specialized industrial domains (e.g., defect detection, safety monitoring)[9].

To overcome this barrier, Synthetic Data Generation[10] has become an active research area. Common methods include using 3D modeling and rendering engines to generate highly realistic images[10] or using Generative Adversarial Networks (GANs) to create new data samples.

A simpler, yet proven highly effective, technique is "Copy-Paste" data augmentation[11]. Research by Ghiasi et al. (2021) demonstrated that even a simple, random "Copy-Paste" operation (pasting an object instance's segmentation mask onto a new background) is an extremely powerful data augmentation method capable of significantly boosting model performance.

2.3 Knowledge Transfer in Industrial Applications

Transfer Learning (TL) is a standard technique for addressing data scarcity in industrial vision[12]. The core idea is to use a model pre-trained on a large, general-purpose dataset (like ImageNet) as a starting point, and then "fine-tune" it on a small, domain-specific target dataset. This allows the model to reuse learned general-purpose, low-level features (like edges and textures) and focus on learning task-specific, high-level features.

Domain Adaptation (DA) is a branch of transfer learning that aims to solve the problem of distribution discrepancy (i.e., "domain gap") between a source domain (e.g., synthetic data) and a target domain (e.g., real data)[13]. DA techniques have been successfully applied to synthetic smoke detection[10] and other fire recognition tasks[13], with the goal of training a model on the source domain that performs well on the target domain.

2.4 Early Warning as Small Object Detection

The industrial safety requirement for early detection[4] technically translates directly into a Small Object Detection (SOD) challenge. SOD is difficult because small objects (like an incipient flame) may only be a few pixels in size, and their weak feature information can be easily diluted or lost entirely during the continuous down-sampling process of a CNN[5].

Current mainstream solutions for SOD primarily focus on network architecture improvements, such as using a Feature Pyramid Network (FPN)[14] to fuse multi-scale features, adding extra detection heads for high-resolution feature maps, or introducing attention mechanisms.

Our work addresses this problem from a complementary angle—the data level. By explicitly controlling the training data distribution, we force the model to pay more attention to small objects, which is critical for applications like UAV aerial imagery or industrial monitoring[9].

2.5 AI in Industrial and Biogas Systems

Artificial intelligence has already demonstrated strong potential in industrial automation and energy systems. For example, YOLO-MSD, developed by Ge et al.[9], is specialized for industrial surface defect detection. This work focuses on solving the multi-scale feature fusion problem for tiny defects, which is inherently similar to the early flame (small object) detection challenge we face.

Furthermore, YOLO-GD, developed by Yue et al.[15], successfully applies the YOLO framework to the specific industrial automation task of robotic grasping.

Recent work has also explored intelligent methods in biogas systems beyond safety monitoring. Geng et al. proposed BiogasNET with BiogasGAN for enhanced biogas production prediction[16], further highlighting the growing role of AI in smart biogas facilities.

3 METHODOLOGY

Bio-FlameNet follows a two-stage training workflow: general feature pre-training on a large public flame dataset, followed by domain-specific adaptation using a synthetic biogas-flame dataset.

3.1 D-Fire General Feature Pre-training

Source Dataset. This phase leverages the public D-Fire dataset[6], a large collection of over 21,000 images and more than 14,000 annotated flame instances, primarily captured in wildland scenes such as forests and grasslands.

Objective. The goal is to pretrain the model on D-Fire to obtain robust weight initialization that captures universal visual fire features—low-level color and texture cues and mid-level morphological patterns.

3.2 Domain-Specific Synthetic Data Generation

This phase constitutes the core of our methodology, addressing both data scarcity and scene specificity through synthetic data generation. We adopt a copy–paste image-fusion pipeline: a curated library of real biogas facility backgrounds (pipes, tanks, valves, complex lighting and reflections) is combined with a library of flame assets (transparent-channel sprites of varying shapes and sizes) to synthesize training images by randomly pasting one or more flame assets onto background scenes.

Key innovation — mathematical size distribution control: to prioritize early detection of small fires, we replace uniform size sampling with a linearly decreasing probability density function, implemented via inverse-transform sampling; this yields a targeted small:medium:large ratio of 7:2:1[5]. The resulting dataset intentionally biases training toward small targets, forcing the model to learn subtle features associated with early incidents.

Multi-level augmentation: background augmentation (brightness, contrast, hue shifts; blur; noise) simulates environmental and camera variations, while flame augmentation (random rotation, color jitter, transparency changes) simulates flame dynamics and occlusion.

3.3 Transfer Learning and Fine-tuning

This phase adapts the general knowledge acquired in Phase A to the domain-specific data synthesized in Phase B. Two parallel training strategies were designed for rigorous comparison.

Baseline (train from scratch): the YOLOv8 model is initialized with random weights and trained solely on the Biogas-Flame Synthetic Dataset for 200 epochs.

Proposed (pretrain + fine-tune): the YOLOv8 model is initialized with D-Fire pretrained weights and fine-tuned on the Biogas-Flame Synthetic Dataset for 150 epochs.

Fine-tuning rationale: fine-tuning uses a substantially smaller learning rate (e.g., 1×10^{-4} vs 1×10^{-3}), since the pretrained model already encodes robust feature extractors; fine-tuning therefore adjusts these features to the target domain while mitigating catastrophic forgetting, enabling faster and more stable convergence with fewer epochs and lower computational cost.

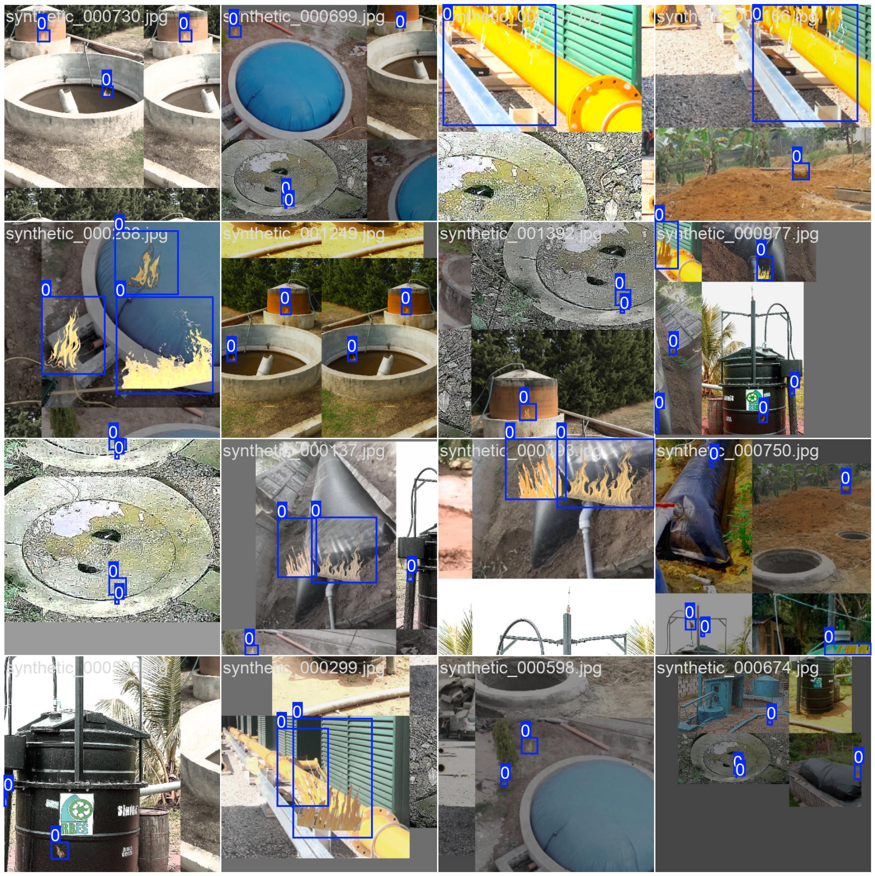


Fig. 1. Samples from the generated synthetic dataset. Flame assets are fused onto real biogas facility backgrounds using the proposed copy-paste pipeline.

4 EVALUATION

4.1 Experimental Setup

Models

We evaluated four core variants from the YOLOv8 series [17]: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), and YOLOv8l (large). These models span a spectrum from lightweight, high-speed configurations to high-capacity, high-accuracy architectures, enabling analysis of strategy effectiveness under varying computational constraints.

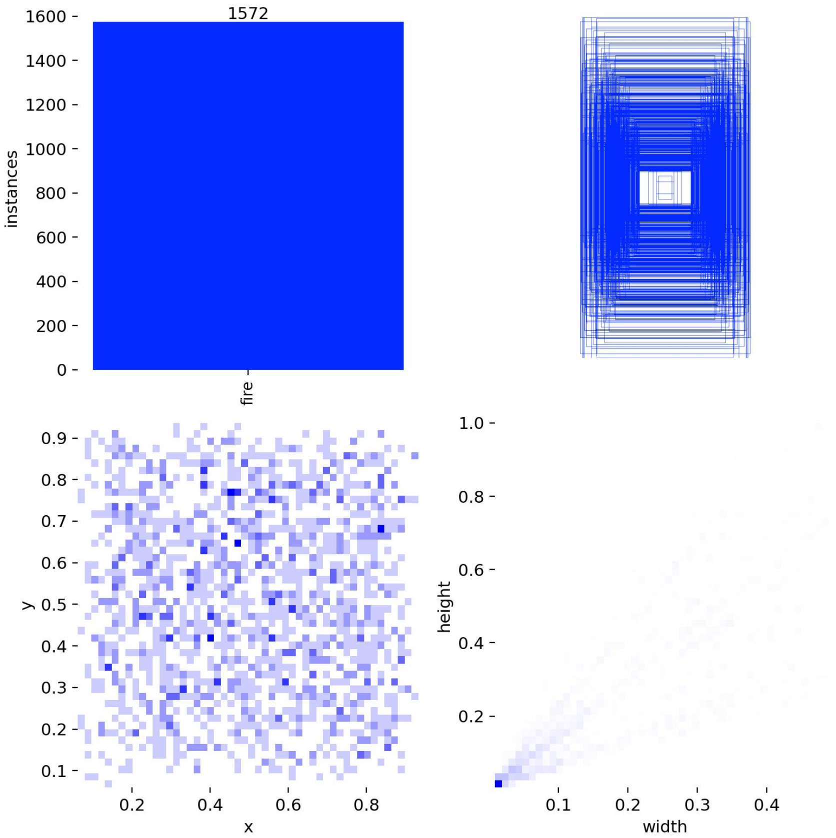


Fig. 2. Visualization of label distribution. The dataset follows a targeted 7:2:1 ratio (Small:Medium:Large) to prioritize early flame detection.

Dataset

All experiments were conducted on the *biogas_synthetic_merged* dataset (see Section III-B), which follows a 70:20:10 distribution of small, medium, and large flame instances, respectively.

Training Strategies

We compare two training regimes as defined in Section III-C:

- **Train from Scratch:** 200 epochs with randomly initialized YOLOv8 weights.
- **Pre-train + Fine-Tune:** 150 epochs using D-Fire pretrained weights from Phase 1.

Evaluation Metrics

To evaluate detection performance we compute counts from the prediction confusion matrix and derive standard detection metrics. From each test set we extract True Positives (TP), False Positives (FP), False Negatives (FN), True Negatives (TN):

- **Precision:** Measures the proportion of correct positive detections; critical for minimizing false alarms in industrial settings.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

- **Recall:** Measures the proportion of actual positives correctly identified; essential for ensuring safety in fire detection.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

- **Intersection over Union (IoU):** Measures overlap between a predicted box B_p and a ground-truth box B_{gt}:

$$\text{IoU}(B_p, B_{gt}) = \frac{\text{area}(B_p \cap B_{gt})}{\text{area}(B_p \cup B_{gt})}. \tag{3}$$

  IoU is used to decide whether a prediction matches a ground truth.

- **mAP@50:** Mean Average Precision at an IoU threshold of 0.5, following the PASCAL VOC standard.

$$\text{mAP@50} \;=\; \frac{1}{N}\sum_{i=1}^{N} \text{AP}_i(0.50). \tag{4}$$

- **mAP@50–95:** COCO-style mean Average Precision averaged over IoU thresholds from 0.5 to 0.95 in 0.05 increments, requiring higher localization precision.

$$\text{mAP}[0.50:0.95] = \frac{1}{10}\sum_{i=0}^{9} \text{mAP}_{0.50+0.05i} \tag{5}$$

- **Training Time:** Captures convergence speed and computational cost.

4.2 Quantitative Results

The results of all comparative experiments are summarized in Table 1 for direct comparison and in-depth analysis.

Finding 1. Pretraining plus fine tuning substantially reduces convergence time across all evaluated models, with time savings ranging from 25.6% for YOLOv8n to 67.0% for YOLOv8l. This finding supports the transfer-learning motivation: D-Fire pretrained weights provide a superior initialization that lets models concentrate on domain adaptation rather than relearning basic flame features. Our experiments confirm consistent and significant improvements in training efficiency.

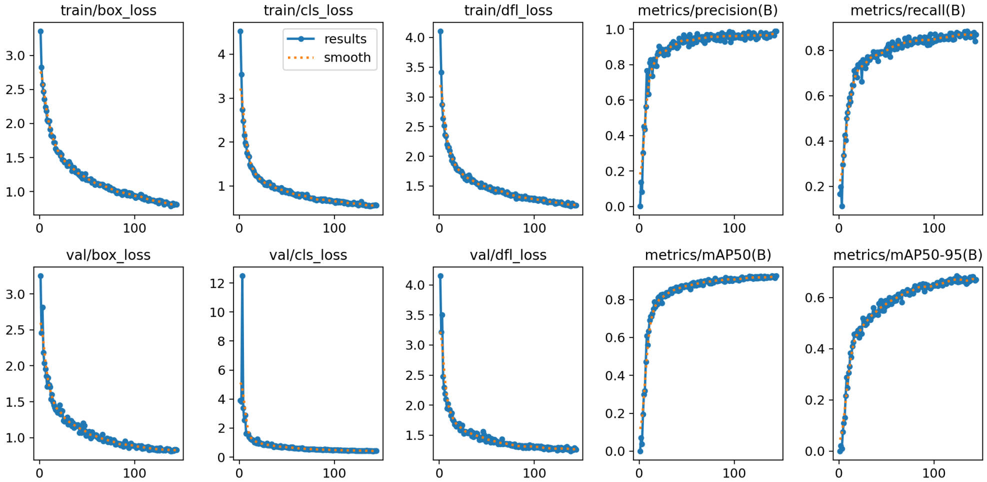


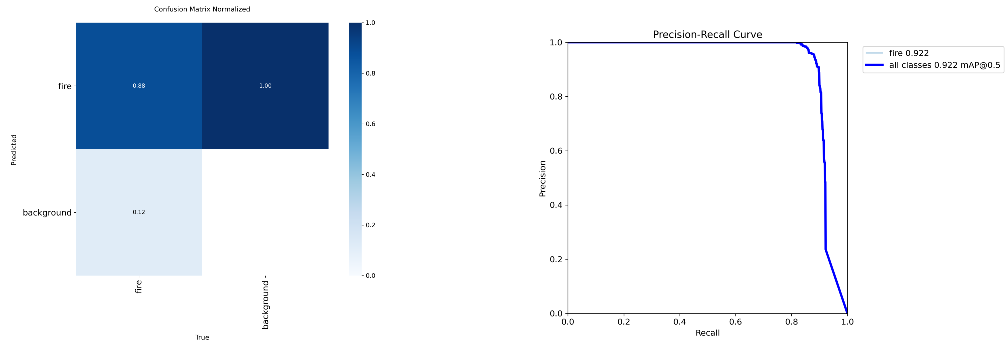
Fig. 3. Training metrics comparison. The curves demonstrate that the proposed "Pre-train + Fine-Tune" strategy converges significantly faster than training from scratch.

Table 1: Performance Comparison of YOLOv8 Models Under Two Training Strategies

Model	Strategy	mAP@50	mAP@50-95	Precision	Recall	Time (min)
YOLOv8n	Train from Scratch	0.9173	0.6793	0.9542	0.8747	9.0
	Ours (Pre-train+FT)	**0.9220**	**0.6969**	**0.9688**	**0.8859**	**6.7**
YOLOv8s	Train from Scratch	0.9224	0.6852	0.9579	0.8814	9.4
	Ours (Pre-train+FT)	0.9225	0.6763	**0.9847**	0.8702	**5.0**
YOLOv8m	Train from Scratch	0.9311	**0.7058**	0.9618	0.8926	24.1
	Ours (Pre-train+FT)	**0.9368**	0.6992	**0.9784**	**0.9133**	**9.0**
YOLOv8l	Train from Scratch	**0.9381**	**0.7133**	0.9690	**0.8881**	37.3
	Ours (Pre-train+FT)	0.9374	0.6957	**0.9759**	0.8837	**12.3**

Finding 2. Under the Pre-train + Fine-Tune regime, YOLOv8n achieves consistent improvements across all metrics: mAP@50 +0.50%, mAP@50–95 +2.59%, Precision +1.53%, and Recall +1.28%. For resource-constrained edge deployment, the two-stage strategy effectively transfers general flame priors and enhances small target detection.

Finding 3. Unlike the lightweight YOLOv8n, higher-capacity models (YOLOv8s, YOLOv8m, YOLOv8l) exhibit a consistent trade-off under the Pre-train + Fine-Tune regime: mAP@50 and Precision improve, while mAP@50–95 decreases. This pattern indicates a domain gap between the D-Fire source domain and the target industrial biogas scenes. Crucially, Precision increases for all fine-tuned high-capacity models, which is often preferable in industrial safety contexts because it reduces false alarms and costly interventions.



(a) Normalized Confusion Matrix

(b) Precision-Recall Curve

Fig. 4. Detailed performance analysis. (a) Normalized confusion matrix confirming the low false positive rate. (b) Precision-Recall curve demonstrating model robustness.

4.3 Qualitative Analysis

To more intuitively demonstrate the findings above, we conducted a series of qualitative analyses.

5 CONCLUSION

This paper presents Bio-FlameNet, a data-centric framework tailored for early flame detection in the high-risk, data-scarce environment of biogas digesters. We demonstrate that a systematic synthetic data pipeline, incorporating mathematical size distribution control to enforce a 7:2:1 flame size ratio, effectively addresses the Small Object Detection (SOD) challenge at the data source level. Comprehensive experiments confirm that our two-stage transfer learning strategy—combining D-Fire pre-training with domain-specific fine-tuning—substantially improves training efficiency, reducing convergence time by up to 67%. Furthermore, this approach significantly enhances detection performance on lightweight models (e.g., YOLOv8n) by transferring general fire priors, making it a viable solution for resource-constrained edge deployment. For high-capacity models, the strategy acts as a domain regularizer, effectively improving Precision to minimize false alarms in safety-critical industrial settings.

Future research will focus on bridging the remaining sim-to-real domain gap. First, we will explore unsupervised domain adaptation (UDA) techniques[13] leveraging unlabeled surveillance footage to align feature distributions between synthetic and real domains. Second, we plan to enhance data realism by implementing background-aware copy-paste augmentation[11] and incorporating a hybrid dataset of real-world hard negatives (e.g., metallic reflections) to further improve discriminative capacity. Third, we aim to develop a multi-modal fusion pipeline integrating thermal infrared (IR) sensing to improve robustness under challenging lighting conditions. Finally, the refined models will be quantized and deployed on edge devices for long-term field validation in operational biogas facilities.

Acknowledgments

Part of this work are supported by Japan Science and Technology Agency (JST), Africa-Japan Collaborative Research (AJ-CORE).

Declaration on Generative AI

During the writing of this article, the author used Gemini for grammar and spell checking, as well as language polishing.

References

[1] Westenbroek PA, Martin I J. Anaerobic Digesters and Biogas Safety. Livestock and Poultry Environmental Learning Community (LPELC). 2019.

[2] U S Environmental Protection Agency. Safety Practices for On-Farm Anaerobic Digestion Systems. EPA; 2020. EPA-430-B-20-003.

[3] Hassan A, Audu AI. Traditional sensor-based and computer vision-based fire detection systems: A review. International Journal of Scientific & Engineering Research. 2022;13(9).

[4] de Venancio PVA, Lisboa AC, Barbosa AV. An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices. Neural Computing and Applications. 2022;34:15349-68.

[5] Wahyudi D, Soesanti I, Nugroho HA. Toward detection of small objects using deep learning methods: A review. In: 14th International Conference on Information Technology and Electrical Engineering (ICITEE); 2022. p. 314-9.

[6] Elhanashi A, Essahraui S, Dini P, Saponara S. Early fire and smoke detection using deep learning: A comprehensive review of models, datasets, and challenges. Applied Sciences. 2025;15(18):10255.

[7] Chen T, Liu Y, Wu B. Flame region detection using color and motion features in video sequences. In: Chinese Control and Decision Conference (CCDC); 2014. p. 2501-5.

[8] Redmon J, Divvala S, Girshick R, Farhadi A. You Only Look Once: Unified, Real-Time Object Detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 779-88.

[9] Ge Y, Li Z, Meng L. YOLO-MSD: A robust industrial surface defect detection model via multi-scale feature fusion. Applied Intelligence. 2025;55:840.

[10] Xu G, Zhang Y, Zhang Q, Lin G, Wang J. Deep domain adaptation based video smoke detection using synthetic smoke images. Fire Safety Journal. 2017;93:53-9.

[11] Ghiasi G, Cui Y, Srinivas A, Qian R, Lin TY, Cubuk ED, et al. Simple Copy-Paste Is a Strong Data Augmentation Method for Instance Segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2021. p. 2918-28.

[12] Umair M, Foo YL. Industrial safety helmet detection using single shot detectors models and transfer learning. In: Multimedia University Engineering Conference (MECON 2022). Atlantis Press; 2022. p. 390-400.

[13] Yan Z, Wang L, Qin K, Zhou F, Ouyang J, Wang T, et al. Unsupervised domain adaptation for forest fire recognition using transferable knowledge from public datasets. Forests. 2023;14(1):52.

[14] Lin TY, Dollar P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2017. p. 2117-25.

[15] Yue X, Li H, Shimizu M, Kawamura S, Meng L. YOLO-GD: A deep learning-based object detection algorithm for empty-dish recycling robots. Machines. 2022;10(5):294.

[16] Geng Y, Wang Z, Deng M, Meng L. Enhanced Biogas Production Prediction Using BiogasNET with BiogasGAN. In: 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 2025. p. 7133-8.

[17] Ultralytics. YOLOv8; 2024. [Online]. Available: https://github.com/ultralytics/ultralytics (accessed May 30, 2026). GitHub repository.