

Article

LGG-YOLO-Seg: A Lightweight Real-Time Instance Segmentation Framework for Low-Grade Graphite Ore Sorting

Jionghui Wang ¹, Yuxing Yu ¹, Guangzheng Yao ¹, Zeyang Qiu ^{2,3}  and Zhaojie Sun ^{4,*}

¹ China Minmetals Corporation (Heilongjiang) Graphite Industry Co., Ltd., Hegang 154100, China; wangjh@minmetals.com (J.W.); yuyux@minmetals.com (Y.Y.); yaogz6@minmetals.com (G.Y.)

² School of Metallurgical Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China; qq2462218642@163.com

³ Ganzhou Chuangzhigu Electronic Intelligent Manufacturing Industry Research Institute, Ganzhou 341600, China

⁴ School of Industrial Software, Jiangxi University of Science and Technology, Nanchang 330013, China

* Correspondence: 15070829913@163.com

Abstract

Real-time instance segmentation of low-grade graphite ores is essential for intelligent mineral-sorting systems, yet existing methods often struggle to balance segmentation accuracy, computational efficiency, and lightweight deployment under irregular boundaries, multi-scale targets, and complex illumination conditions. To address these challenges, this paper proposes LGG-YOLO-Seg, a lightweight and efficient instance segmentation framework based on YOLO11-seg. First, a compact baseline is constructed through depth scaling to reduce redundant computation and model complexity. Then, Linear Deformable Convolution (LDConv) is embedded into the backbone to enhance adaptive feature extraction for irregular ore boundaries. Meanwhile, the feature extraction module is redesigned as C3k2-PKI using multi-kernel and depthwise separable convolutions to improve multi-scale representation capability. In addition, a lightweight Segment-EfficientHead is introduced to reduce computational overhead while maintaining segmentation accuracy. Experiments conducted on a self-constructed graphite ore dataset demonstrate that LGG-YOLO-Seg improves mAP_{50_{box}} and mAP_{50_{mask}} by 0.9% and 0.4%, respectively, compared with YOLO11n-seg while reducing FLOPs, parameters, and model size by 28.8%, 55.6%, and 55.1%, respectively, with an 11.0% increase in inference speed, showing strong potential for real-time industrial sorting and edge deployment.

Keywords: instance segmentation; LGG-YOLO-Seg; graphite ore sorting; deformable convolution; lightweight model; edge deployment



Academic Editor: Silvia Liberata Ullo

Received: 28 May 2026

Revised: 24 June 2026

Accepted: 7 July 2026

Published: 9 July 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.

1. Introduction

Graphite ore is a critical strategic resource supporting the development of high-tech industries such as new energy and advanced materials. Owing to its multifunctionality and unique properties, graphite has become an indispensable material across numerous industrial and high-end technological applications [1,2]. Notably, graphite maintains excellent hardness and strength even at extreme temperatures up to 3600 °C [3]. Moreover, its characteristic layered crystal structure endows it with a range of outstanding properties, enabling widespread use in cutting-edge fields including lithium-ion batteries, fuel cells, electronic devices, optical fibers, spintronics, refractory materials, consumer electronics, and new energy electric vehicles [4,5]. In particular, with the rapid growth of the electric

vehicle sector in recent years, graphite, as the core anode material in lithium-ion batteries, directly influences key performance metrics such as cycle life, energy density, and power density, and indirectly affects battery design and overall performance [6].

The grade of graphite ore refers to the mass percentage of valuable carbon components and serves as a key indicator for evaluating fixed carbon content. This metric not only reflects the economic value of mineral resources but also provides an essential basis for assessing mining feasibility and guiding operational decision-making [7]. Low-grade graphite ores are typically characterized by low carbon content, abundant impurities, and complex structures [8], resulting in limited overall utilization efficiency. Inadequate screening and identification during production may lead to the mixing of low-grade ores into downstream processes, which not only reduces product stability but also significantly increases energy consumption and equipment wear, thereby raising production costs and causing resource waste. Moreover, inefficient mineral sorting can increase tailings' discharge, posing adverse impacts on the ecological environment and green mining development [9]. Therefore, developing an accurate and efficient method for low-grade graphite ore identification, enabling fast and reliable intelligent discrimination at the front end of sorting, is of great significance for improving resource utilization, reducing energy consumption and environmental burden, and promoting the intelligent transformation of graphite processing.

Traditional graphite ore sorting has long relied on a combination of manual experience and physical and chemical methods, including crushing, flotation, and reagent-based treatments to achieve mineral separation [10,11]. Such approaches are often characterized by lengthy workflows, high energy consumption, and strong dependence on operator expertise. In addition, the use of chemical reagents may lead to environmental pollution, making these methods less suitable for modern requirements of green mining and efficient production. Subsequently, machine learning techniques were introduced into ore sorting, where features such as color, texture, and morphology are combined with classification models for automated identification [12], improving automation and processing efficiency to some extent. However, their performance heavily depends on the quality of handcrafted features and tends to be unstable under complex illumination and noisy backgrounds. Furthermore, they are limited in their ability to extract deep discriminative representations from raw data, which restricts their robustness in real-world mining environments [13,14]. In recent years, deep learning has driven the intelligent development of graphite ore sorting. Models based on object detection and image segmentation can learn discriminative features in an end-to-end manner, significantly improving recognition accuracy and robustness [15,16]. Nevertheless, existing deep learning models are often characterized by large parameter sizes and high computational complexity, along with stringent hardware and real-time requirements, which hinder their deployment in on-site mining scenarios and resource-constrained edge devices.

To address the limitations of traditional and machine learning methods in feature representation and environmental adaptability, as well as the difficulty of existing deep learning approaches in balancing segmentation accuracy and efficiency, this paper proposes an efficient instance segmentation method for low-grade graphite ore based on an improved YOLO11-seg framework. The proposed method not only significantly enhances segmentation accuracy but also effectively reduces model complexity and parameter scale, achieving a favorable trade-off between performance and inference efficiency. This provides a practical and effective solution for accurate identification and efficient sorting of low-grade graphite ores. The main contributions of this work are summarized as follows:

- To tackle practical challenges such as high annotation cost, limited data availability, and complex on-site imaging conditions, a representative graphite ore dataset is constructed. Ore samples are collected from real mining environments and their

grade levels are accurately determined through chemical analysis. Images are captured under multiple exposure conditions to simulate realistic illumination variations, and finely annotated to form a dataset containing two grade categories and three exposure settings, providing a reliable foundation for model training and evaluation. This dataset also provides a task-specific basis for evaluating lightweight instance segmentation methods in low-grade graphite ore-sorting scenarios.

- Without altering the overall architecture and global width of the original network, the model depth and maximum channel parameters are adjusted to reduce redundant module stacking in both the backbone and feature fusion layers. Based on this, a more compact YOLO11cp-seg model is developed, which significantly decreases computational cost and parameter size while preserving essential feature representation capability for segmentation tasks.
- To the best of our knowledge, this study is among the first attempts to introduce Linear Deformable Convolution (LDConv) into low-grade graphite ore instance segmentation. Linear Deformable Convolution (LDConv) is introduced into the backbone to replace conventional downsampling convolutions, enabling adaptive receptive fields that dynamically adjust to ore shapes. This design preserves critical texture and boundary information while suppressing background interference, thereby improving multi-scale feature extraction and enhancing robustness and efficiency in complex industrial environments. Compared with standard convolutional downsampling, this design is more suitable for boundary-sensitive feature extraction of irregular ore particles.
- The original C3k2 module is redesigned by embedding a multi-scale PKI structure to form the C3k2-PKI module. This design is intended as an application-specific feature extraction enhancement for low-grade graphite ore images, rather than a completely new multi-kernel modeling paradigm. The module captures features of ores at different scales through multi-level convolutions while employing depthwise separable convolutions to control computational overhead, thereby enhancing information flow and multi-scale feature aggregation. Compared with standard feature extraction modules, C3k2-PKI is better suited to the multi-scale particle distribution and complex surface textures of low-grade graphite ores.
- For industrial segmentation scenarios with limited categories and relatively simple structures, the segmentation head is redesigned as a Segment-EfficientHead. The purpose of this design is to reduce redundant computation in the segmentation head according to the task characteristics of low-grade graphite ore sorting. By optimizing convolutional structures and adopting parameter sharing, the proposed head reduces computational burden and model size while maintaining mask accuracy, enabling faster inference and lower resource consumption suitable for real-time deployment in industrial settings. This task-oriented lightweight head differs from general segmentation heads by reducing redundant computation according to the limited-category characteristics of low-grade graphite ore sorting.

2. Related Work

Traditional ore-sorting methods predominantly rely on manual experience, supplemented by physical separation and chemical processing techniques. Typically, ores are first roughly separated based on visual appearance and empirical judgment through manual sorting, followed by mineral separation using processes such as crushing, flotation, and reagent treatment. Although these conventional approaches can achieve relatively accurate sorting results to some extent, they suffer from several inherent limitations. In particular, the rising cost of manual labor and the prolonged sorting cycle significantly reduce their economic efficiency [17]. Moreover, the heavy dependence on operator expertise intro-

duces strong subjectivity, resulting in limited stability and poor standardization, making quantitative analysis difficult. In addition, traditional graphite ore-grade evaluation primarily depends on physical and chemical analysis methods, including manual sampling and instrumental techniques such as X-ray diffraction (XRD) and LECO carbon–sulfur analyzers [18,19]. For instance, Cui et al. [20] analyzed the physicochemical properties of flake graphite ore using XRD and electron microscopy, providing insights for beneficiation process optimization. Similarly, Sader et al. [21] employed a LECO carbon–sulfur analyzer to combust samples and measure infrared absorption of carbon dioxide, enabling accurate determination of total carbon content in graphite ores. Despite their high accuracy, these methods are characterized by complex procedures, long processing cycles, and strong dependence on labor and equipment, along with low efficiency in data acquisition and organization, making them unsuitable for rapid, large-scale grade evaluation and efficient segmentation tasks.

With the rapid advancement of machine vision and machine learning technologies, ore-sorting methods that integrate mineral image analysis with learning algorithms have gradually matured, emerging as an effective approach to improving sorting efficiency and automation [22]. Early machine vision-based studies primarily relied on handcrafted features such as color, texture, edges, and shape, combined with classifiers including support vector machines, random forests, and shallow neural networks for ore recognition. While these methods can achieve reasonable performance under controlled conditions and simple backgrounds, their effectiveness is highly sensitive to feature design and image acquisition environments. In complex scenarios involving surface reflections, occlusion, adhesion, and subtle grade differences, they often suffer from insufficient feature representation and limited generalization capability. For example, Shatwell et al. [23] proposed a real-time ore-sorting method based on color and texture analysis, integrating image processing features with machine learning models to enable rapid classification of gold and silver ore grades. Their study demonstrated the practical value of traditional visual features in real-time sorting while also highlighting limitations in scene adaptability and feature robustness.

With the continuous advancement of deep learning, researchers have increasingly leveraged convolutional neural networks (CNNs) to automatically learn more discriminative deep features from mineral images, thereby significantly reducing the reliance on handcrafted feature engineering. Compared with traditional machine learning approaches, deep learning models can more effectively capture ore texture, color distribution, and local structural variations under complex backgrounds, demonstrating superior robustness and adaptability in ore image classification and grade recognition tasks. For instance, Liu et al. [24] systematically investigated the impact of network depth, model architecture, and dataset scale on ore image classification performance, highlighting the strong application potential of lightweight deep models in ore-sorting scenarios. Furthermore, they summarized recent advances in deep learning for digital mineral image classification, indicating a paradigm shift from shallow models relying on manual expertise toward end-to-end feature learning frameworks, which exhibit clear advantages in multi-class mineral recognition, complex texture modeling, and industrial automation analysis [25].

Building upon these developments, mineral image analysis has gradually evolved from image-level classification to pixel-level segmentation. Compared with classification and object detection, image segmentation provides more precise boundary delineation and region-level representation, making it more suitable for industrial scenarios characterized by complex particle contours, severe adhesion, and significant scale variations. Moreover, segmentation offers fine-grained data support for downstream tasks such as particle size analysis and automated process control. For example, Ferreira et al. [26] employed Mask R-CNN to perform instance segmentation of quartz particles in iron ore

microscopic images, achieving particle-level precise recognition. Similarly, Tang et al. [27] proposed an improved YOLOv8n-based method for mineral recognition and segmentation, demonstrating promising performance in handling fine-grained targets.

From the perspective of general computer vision development, models such as U-Net, Mask R-CNN, and YOLACT have provided important references for industrial mineral segmentation. Ronneberger et al. [28] proposed U-Net, which enhances the fusion of fine-grained details and semantic information through an encoder–decoder architecture. He et al. [29] introduced Mask R-CNN, achieving high-precision instance segmentation by incorporating an additional mask branch. Bolya et al. [30] proposed YOLACT, which balances segmentation accuracy and real-time performance. However, when directly applied to low-grade graphite ore scenarios, these models still suffer from limitations such as large parameter sizes, high computational overhead, and insufficient adaptability to complex boundaries and multi-scale targets.

In recent years, deep learning research on ore recognition and grade detection has increasingly focused on lightweight design and real-time deployment. For example, Xiang et al. [31] proposed a graphite ore-grade recognition method that combines Faster R-CNN with an attention mechanism, improving the representation of graphite texture features; however, the two-stage architecture still limits real-time performance. Qiu et al. [32] further proposed Stellar-YOLO for graphite ore-grade detection, achieving a favorable balance between accuracy and lightweight design. Nevertheless, this method mainly targets object-level recognition and lacks dedicated design for instance-level mask generation and fine-grained region representation. Moreover, its focus remains on detection performance optimization, with relatively limited exploration of precise segmentation for complex boundary targets.

Overall, existing studies have laid a solid foundation for intelligent ore recognition and automated sorting, yet several challenges remain. Traditional vision-based methods heavily depend on handcrafted features and environmental conditions, while a considerable portion of deep learning research still focuses on classification or detection tasks. Meanwhile, general-purpose segmentation models often struggle to simultaneously balance accuracy, lightweight design, and real-time performance. Therefore, considering the challenges of complex illumination, irregular boundaries, and multi-scale distribution in low-grade graphite ores, it is of great significance to develop efficient and lightweight instance segmentation methods.

3. Methods

3.1. Original YOLO11-Seg Model

The YOLO series formulates object detection as a regression problem, enabling end-to-end optimization and real-time inference [33]. In 2024, Ultralytics released the YOLO11-seg algorithm, which includes five variants with different parameter scales (n, s, m, l, and x) [34]. As one of the state-of-the-art instance segmentation models, its architecture mainly consists of three components: a backbone, a neck, and a head. The overall structure and detailed configurations are illustrated in Figure 1.

Compared with its predecessors, YOLO11 [35] introduces systematic improvements at the architectural level. First, the backbone adopts a newly designed C3k2 module to replace the original C2f module. By integrating variable convolution kernels and channel separation strategies, this module significantly enhances deep feature extraction capability in complex scenarios while reducing the number of parameters. Second, following the Spatial Pyramid Pooling Fast (SPPF) layer, a C2PSA module equipped with a multi-head attention mechanism is introduced, which effectively strengthens multi-scale feature perception, fusion, and robustness to occlusion. Finally, the classification branch in the

detection head incorporates depthwise separable convolution (DWConv), substantially reducing computational overhead and accelerating inference.

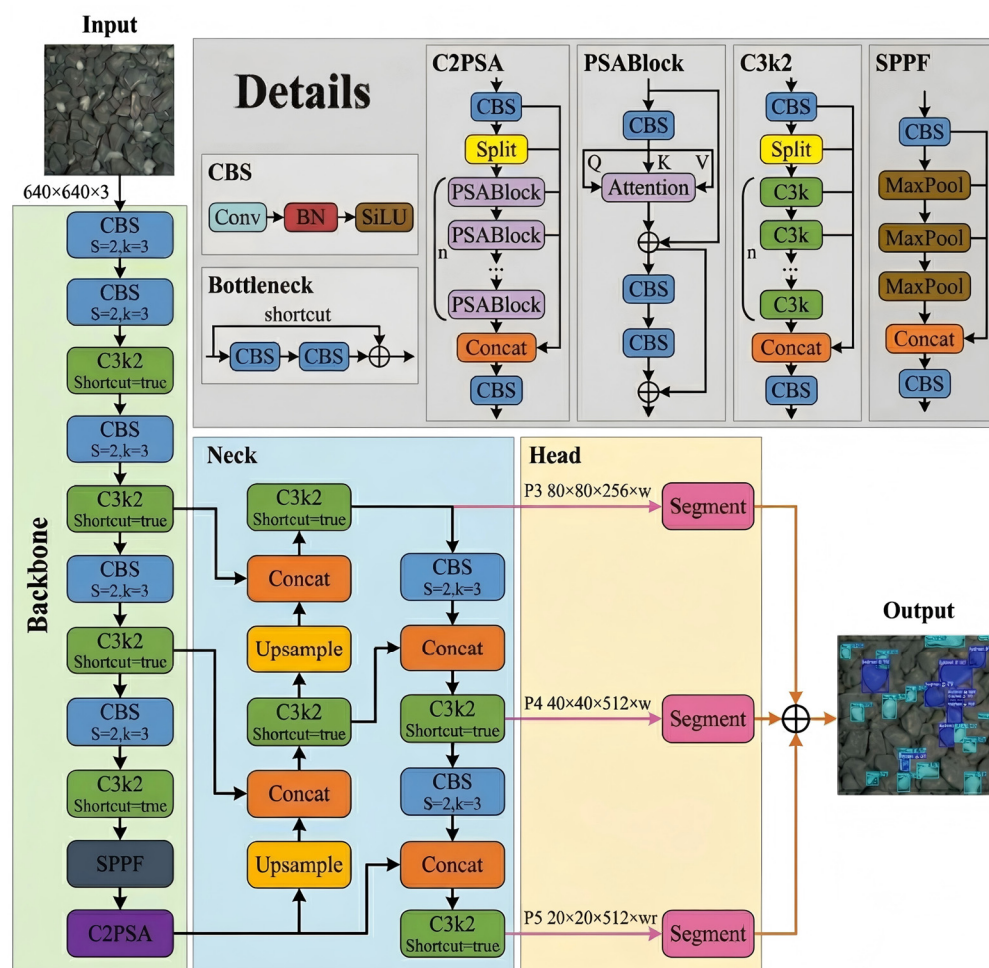


Figure 1. Architecture of the YOLO11-seg network.

Considering the limited computational resources of edge devices in industrial sorting scenarios of low-grade graphite ores, YOLO11n-seg is selected as the baseline model due to its favorable balance between accuracy and lightweight deployment. Based on this baseline, further improvements in the backbone and feature extraction architecture are conducted to overcome performance bottlenecks and achieve efficient and intelligent sorting of graphite ores.

3.2. Improved LGG-YOLO-Seg Model

To address the instance segmentation task of low-grade graphite ores in industrial environments, this study proposes an improved LGG-YOLO-Seg model based on YOLO11-seg, while also constructing YOLO11cp-seg as a lightweight transitional baseline. YOLO11cp-seg retains the overall architecture and global width of the original network, but adjusts network depth and optimizes module stacking, substantially reducing redundant computations in both the backbone and feature fusion layers. This design preserves essential feature representation capability while achieving lower parameter count and computational overhead, providing a compact starting point for further enhancements.

Within the backbone, LGG-YOLO-Seg replaces standard downsampling convolutions with Linear Deformable Convolutions (LDConv), enabling the convolutional kernels to adaptively adjust receptive fields according to ore morphology and scale. This modification enhances feature capture for irregular boundaries and complex textures, while avoiding

the parameter and computational growth associated with conventional large receptive fields. Furthermore, the original C3k2 module is redesigned as C3k2-PKI, which combines multi-scale convolutional kernels with depthwise separable convolutions to expand feature propagation paths and strengthen multi-scale information aggregation. This ensures robust feature representation for both fine-grained particles and large ore blocks, improving segmentation accuracy while controlling computational complexity.

Finally, for the segmentation head, LGG-YOLO-Seg introduces the Segment-EfficientHead module tailored for industrial scenarios with few categories and relatively simple structures. By optimizing convolutional structures and adopting weight-sharing strategies, this module reduces redundant computation, generates masks efficiently with precise boundaries, and maintains high-confidence predictions, ensuring robust segmentation under varying ore layouts and illumination conditions. The overall architecture of LGG-YOLO-Seg is illustrated in Figure 2.

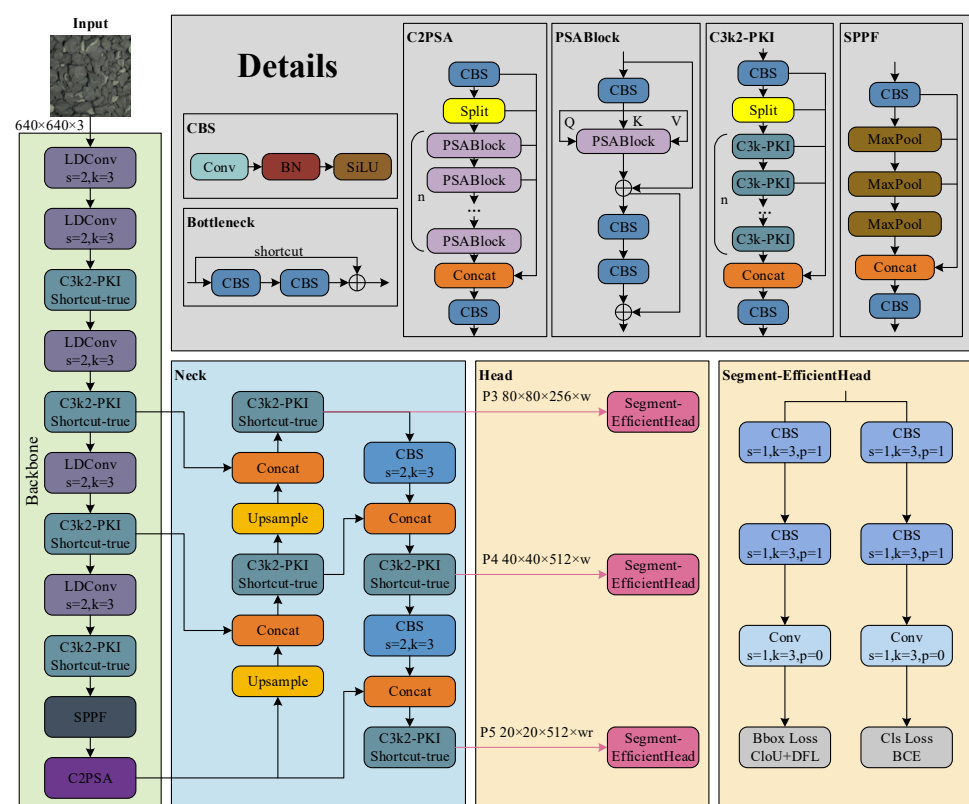


Figure 2. Architecture of the LGG-YOLO-Seg network.

It should be clarified that the proposed LGG-YOLO-Seg is not intended to claim that all its basic components are newly invented from scratch. Several design ideas, such as lightweight scaling, deformable convolution, multi-kernel feature extraction, and efficient prediction heads, are inspired by previous studies. In this work, the novelty mainly lies in their task-oriented modification, coordinated integration, and validation for low-grade graphite ore instance segmentation. Specifically, LDConv is adopted and embedded into the YOLO11-seg backbone to improve boundary-adaptive feature extraction for irregular ore particles. For the C3k2-PKI module, the original C3k2 structure in YOLO11 is retained as the basic feature-flow framework, while the internal bottleneck is modified by incorporating a PKI-based multi-kernel feature extraction structure and depthwise separable convolutions. This modification is designed to enhance multi-scale representation for ore particles with different sizes and complex surface textures while controlling computational overhead. For Segment-EfficientHead, the novelty does not lie in proposing a completely

new segmentation head paradigm, but in redesigning the head according to the limited-category and real-time deployment characteristics of low-grade graphite ore sorting. By simplifying redundant convolutional operations and introducing parameter-sharing strategies, the head reduces computational cost while maintaining mask prediction capability. Therefore, the main contribution of this work is an application-specific lightweight instance segmentation framework that balances segmentation accuracy, model complexity, and inference speed for industrial graphite ore sorting.

3.3. YOLO11cp-Seg

The YOLO11-seg series enables flexible scaling of network size (e.g., n, s, m, l, and x) by introducing three key parameters: depth, width, and max_channels, allowing for adaptation to diverse application requirements [36]. Specifically, the depth parameter controls the number of repeated feature extraction modules (e.g., C3k2), thereby determining the overall network depth. The width parameter adjusts the number of channels in each layer, influencing the network capacity. In addition, max_channels sets an upper bound on the output channels of each layer, preventing excessive model size, memory consumption, and computational burden, thus achieving a balance between performance and efficiency.

Although deeper architectures can enhance the ability to learn complex semantic features, they inevitably incur substantial computational cost and memory overhead. Considering that this study focuses on instance segmentation of low-grade graphite ores, where the number of target categories in practical industrial scenarios is relatively limited, even the smallest baseline model in this series, YOLO11n-seg, still exhibits noticeable structural and channel redundancy [37]. To further alleviate the performance bottleneck of edge devices and improve inference efficiency, a global compression strategy is adopted to develop a shallower and more compact lightweight baseline variant, termed YOLO11cp-seg.

Specifically, while preserving the original network architecture and global width, the depth factor is reduced to 0.25 and the max_channels parameter is set to 512. This targeted adjustment effectively decreases the number of repeated feature extraction modules in both the backbone and neck, thereby eliminating redundant operations during feature extraction. As a result, the proposed YOLO11cp-seg significantly reduces network depth, parameter count, and floating-point operations (FLOPs), establishing a lightweight foundation for subsequent construction of an efficient industrial-grade intelligent sorting model. A comparison of parameters and FLOPs under different scaling configurations of YOLO11-seg is presented in Table 1.

Table 1. Comparison of scaling factors in YOLO11-seg models.

Model	Depth	Width	Max_Channels	Params (M)	FLOPs (G)
YOLO11m-seg	0.5	1	512	22.36	123.6
YOLO11s-seg	0.5	0.5	1024	10.08	35.6
YOLO11n-seg	0.5	0.25	1024	2.84	10.4
YOLO11cp-seg	0.25	0.25	512	1.71	9.3

3.4. Linear Deformable Convolution (LDConv)

In the instance segmentation of low-grade graphite ores, target regions typically exhibit highly irregular shapes and complex surface structures. Ore particles vary significantly in size, boundary contours, and grayscale distribution, often accompanied by adhesion, occlusion, and background interference. Standard convolution, relying on fixed-shape kernels (e.g., 3×3 or 5×5), has a receptive field limited to local neighborhoods, and cannot adaptively adjust to the actual geometric structure of ore particles [38]. Consequently, when processing irregular boundaries or targets with varying scales, standard convolution

tends to incorporate redundant background information, limiting its capacity to represent complex and variable structural features. Directly increasing kernel size to enlarge the receptive field leads to a quadratic growth in model parameters and computational complexity, which is unacceptable for edge devices requiring strict efficiency and memory constraints in industrial real-time sorting systems.

To address these limitations and overcome edge-device performance bottlenecks, we introduce a structurally flexible and computationally efficient Linear Deformable Convolution (LDConv) [39] into the YOLO11n-seg backbone, replacing part of the standard convolution operations. The core of LDConv lies in generating initial sampling coordinates based on a regular grid or prior knowledge and predicting offsets through a lightweight convolutional branch, thereby dynamically adjusting each sampling location. This adaptive sampling mechanism allows the convolution kernel to adjust its receptive field according to the geometric shape of the target, enabling more precise focus on the ore body and effective coverage of multi-scale and irregular targets. This adaptive modeling suppresses background noise and enhances feature extraction flexibility and accuracy for irregular shapes and fine-grained textures. Importantly, LDConv achieves highly efficient parameter utilization: the quadratic parameter growth observed in standard or traditional deformable convolutions is converted into linear growth [40]. The structure of the LDConv module is illustrated in Figure 3.

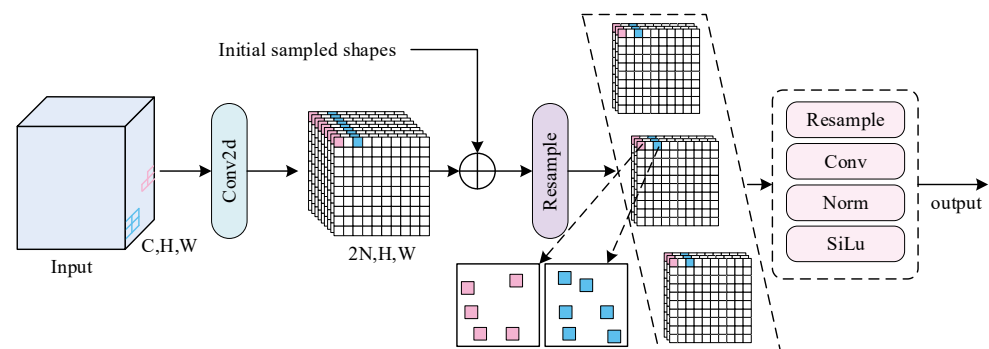


Figure 3. The structure of the LDConv module.

Before constructing the adaptive convolution kernel, the convolution center P_0 is initialized. The weighted aggregation of input features within the local neighborhood is expressed as follows:

$$\text{Conv}(P_0) = \sum_{P_N \in R} w_{PN} \times x(P_0 + P_N). \quad (1)$$

where R represents the set of sampling positions covered by the kernel, w_{PN} is the corresponding weight, and $x(P_0 + P_N)$ denotes the input feature at the relative position $P_0 + P_N$. For an input feature map of dimensions (C, H, W) , a lightweight convolution is employed to learn the spatial offsets, producing a tensor of size $(2N, H, W)$. These offsets are added to the original sampling coordinates $P_0 + P_N$, generating new sampling locations that form an adaptive, irregular spatial pattern. The module then resamples the feature map via interpolation to obtain richer and more discriminative representations. To further improve efficiency and reduce memory access overhead, LDConv fuses resampling, convolution, and normalization operations, minimizing intermediate feature storage and transmission, thereby enhancing computational compactness and execution efficiency.

By integrating LDConv into the backbone, the model achieves a balanced trade-off between computational cost and feature representation, enhances multi-scale feature modeling, reduces overall parameter volume, and provides critical technical support for lightweight deployment in resource-constrained industrial environments.

3.5. C3k2-PKI Module

Compared with earlier feature extraction structures such as C3 and C2f, C3k2 introduces a more flexible feature transmission mechanism that effectively expands information flow paths, thereby enhancing feature diversity and representation capability [41]. Meanwhile, its optimized branch design improves feature extraction efficiency while maintaining fast inference. However, C3k2 still relies on convolution operations with fixed receptive fields [42], which limits its adaptability in low-grade graphite ore instance segmentation. In practical industrial scenarios, ore particles exhibit extreme scale variation: small targets often present blurred boundaries and complex local textures, whereas large targets contain broader global structural information. A single-scale receptive field is insufficient to simultaneously capture fine-grained details and global context, leading to either local semantic loss or incomplete structural perception.

To address this limitation, we embed a Poly Kernel Inception (PKI) structure into C3k2, replacing the original bottleneck and forming the proposed C3k2-PKI module (Figure 4), while preserving the lightweight property of the backbone.

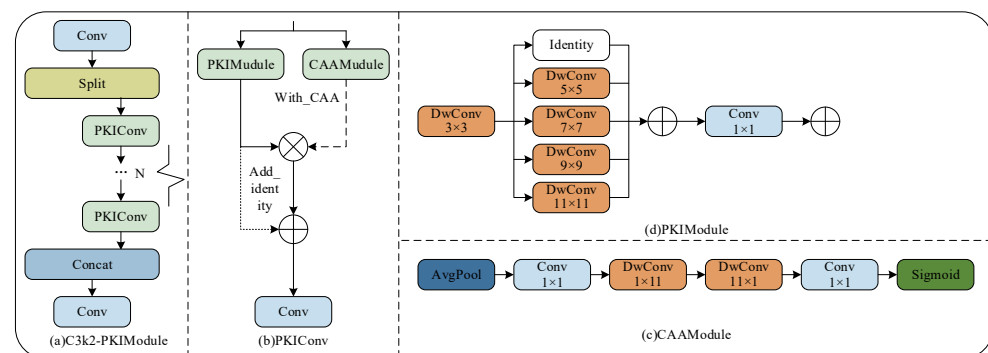


Figure 4. Structure Diagram of the C3K2-PKI Module.

The core design integrates non-dilated multi-scale convolution with a context-aware attention mechanism. First, a Context Anchor Attention (CAA) module is introduced to enhance responses to small-scale ore targets. Specifically, a 7×7 average pooling is applied for low-pass filtering to extract dominant regional features, followed by cascaded cross-directional one-dimensional depthwise convolutions to model spatial dependencies. The attention map is then generated via a Sigmoid activation:

$$Z = \sigma(\text{Conv}_{1 \times 1}(\text{DWConv}_{11 \times 1}(\text{DWConv}_{1 \times 11}(\text{Conv}_{1 \times 1}(\text{AvgPool}_{7 \times 7}(x)))))). \quad (2)$$

where x denotes the input feature tensor, AvgPool represents the average pooling operation, DWConv denotes depthwise separable convolution, σ is the Sigmoid activation function, and Z is the resulting attention-weighted feature.

After attention refinement, the feature map is fed into the PKI core multi-branch feature extraction network. To mitigate background noise introduced by large convolution kernels, the PKI module discards dilated convolutions and instead adopts parameter-efficient parallel multi-scale depthwise separable convolutions (DWConv). Specifically, a 3×3 depthwise convolution is first applied to encode local low-level features and extract fundamental semantic information. The resulting features are then distributed into parallel branches with different receptive fields, i.e., 5×5 , 7×7 , 9×9 , and 11×11 , to capture multi-level contextual information. This multi-scale parallel process can be formulated as follows:

$$x_1 = \text{DWConv}_{3 \times 3}(x), \quad (3)$$

$$y_i = \text{DWConv}_{s \times s}(x_1), s = (i + 1) \times 2 + 1, i = 1, \dots, 4. \quad (4)$$

where x_1 denotes the initially extracted local feature, and y_i represents the output of each parallel scale branch.

Finally, a 1×1 pointwise convolution is employed to perform efficient cross-channel linear fusion of multi-scale receptive field features, formulated as follows:

$$X = \text{Conv}_{1 \times 1}(\sum_{i=1}^4 y_i + x_1). \quad (5)$$

where X denotes the output multi-scale fused feature of the C3k2-PKI module. Notably, although the proposed module significantly enlarges the effective receptive field and incorporates multi-level convolutional kernels, its design strictly adheres to the paradigm of depthwise separable convolution and pointwise fusion, thereby effectively suppressing the exponential growth of parameters and FLOPs. This lightweight yet expressive design enables LGG-YOLO-Seg to robustly handle complex multi-scale ore structures while maintaining high inference speed, making it well-suited for deployment on resource-constrained edge devices.

3.6. Segment-EfficientHead

Unlike general segmentation tasks involving complex backgrounds and numerous categories, this study focuses on low-grade graphite ore segmentation in a single industrial environment, where the number of target categories is extremely limited. The original YOLO11-seg head exhibits structural redundancy in generating bounding boxes and class predictions. To better adapt to this scenario, we redesign the segmentation interface and propose a simplified prediction module termed Segment-EfficientHead.

While strictly preserving the decoupled design between localization and classification, the proposed head removes depthwise separable convolution (DWConv) from the classification branch, thereby streamlining the head structure. This design enables the network to learn task-specific representations for localization, classification, and segmentation independently, effectively reducing feature competition and improving detection accuracy. Meanwhile, eliminating redundant convolution operations significantly reduces computational cost and accelerates inference, making the model more suitable for deployment on resource-constrained edge devices.

The Segment-EfficientHead adopts a dual-branch architecture to efficiently perform object detection and instance segmentation in parallel (Figure 5). Specifically, the detection branch predicts class scores, bounding box offsets, and a set of mask coefficients, which are used to compute classification and localization losses. In parallel, the segmentation branch generates a set of shared prototype masks at the global feature level [43]. During post-processing, non-maximum suppression (NMS) is first applied to remove redundant detections, followed by a linear combination of the predicted mask coefficients and corresponding prototype masks via matrix multiplication to produce instance-specific coarse masks. The segmentation process is optimized using the binary cross-entropy (BCE) loss. Finally, the coarse masks are refined through cropping and thresholding to obtain high-quality instance segmentation results.

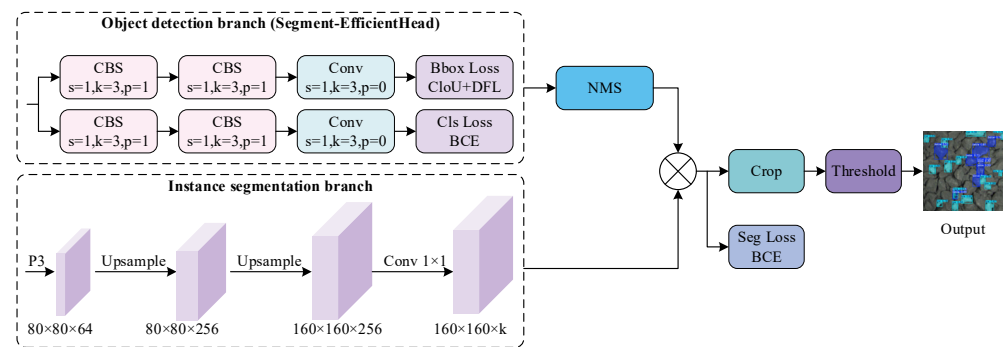


Figure 5. Structure of the proposed Segment-EfficientHead.

4. Data and Experimental Preparation

4.1. Data Collection and Preprocessing

The graphite ore samples used to construct the instance segmentation dataset were collected from the mining area of China Minmetals Group (Heilongjiang) Graphite Industry Co., Ltd., located in Hegang City, Heilongjiang Province, China. This study focused on medium-to-coarse ore fragments (2–15 cm) that were randomly sampled from conveyor belts following secondary crushing. To acquire high-quality visual data, images were captured under the actual visible-light conditions of the beneficiation plant using a Hikrobot MV-CA050-12GC industrial camera (Hikrobot, Hangzhou, China), providing a raw resolution of 3744×3744 pixels. Considering the computational constraints of edge devices and the standardized input requirements of deep learning models, the raw images were cropped and resized to 1024×1024 pixels and stored in JPG format. The collected samples exhibit pronounced compositional differences and clear texture features, making them well-suited for feature extraction by computer vision algorithms.

To address texture degradation caused by high-speed conveyor motion, illumination fluctuations, and ore variability in industrial dynamic sorting, a multi-exposure image acquisition strategy was designed to enhance dataset diversity and surface detail representation. Specifically, three exposure settings were employed: a short exposure of 400 μ s to suppress motion blur under high-speed conditions; a medium exposure of 600 μ s to capture balanced global texture under normal lighting; and a long exposure of 800 μ s to enhance subtle details in low-reflectance or low-light scenarios. This multi-exposure scheme effectively improves feature robustness and segmentation stability in complex industrial environments. A total of 2128 valid images were collected, including 623 short-exposure, 796 medium-exposure, and 709 long-exposure samples. Representative examples under different exposure conditions are shown in Figure 6.

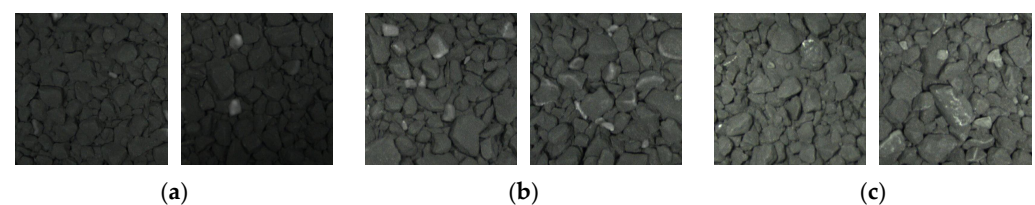


Figure 6. Representative graphite ore images under different exposure modes. (a) short exposure; (b) medium exposure; (c) long exposure.

In graphite ore-grade identification, internal grade is strongly correlated with surface visual characteristics. Specifically, high-grade ores, characterized by low impurity content and strong light absorption, typically exhibit darker appearances. In contrast, low-grade ores often contain light-colored gangue minerals such as quartz and silicates, resulting in

higher surface reflectance and brighter visual appearance. To reduce the subjectivity of category assignment, the grade labels used in this study were not determined solely based on visual appearance. Instead, with the assistance of professional laboratory technicians, the carbon content of ore samples was accurately measured using chemical analysis methods, and the corresponding grade thresholds were determined according to the measured carbon content and expert evaluation. In this study, samples with a carbon content of 0–10% were defined as low-grade graphite ores. To meet fine-grained industrial sorting requirements, these low-grade samples were further divided into two subclasses: the “below” category, corresponding to samples with a carbon content of 0–5%, and the “low” category, corresponding to samples with a carbon content of 5–10%.

During the annotation process, the grade category of each ore sample was first determined based on the chemical analysis results and expert confirmation. Then, instance-level annotations of the corresponding ore particles were manually completed by the author using the Labellmg tool [44] according to the confirmed grade labels. Therefore, the annotation process was guided by objective carbon-content thresholds rather than subjective visual judgment alone. To ensure annotation quality, all annotations were checked after labeling to verify the consistency between the chemical grade information, expert assessment, and instance-level image annotations. For ore particles with ambiguous boundaries, adhesion, or partial occlusion, the annotation was performed along the visible contour of each instance as consistently as possible. When the boundary or category of an instance could not be reliably determined, the sample was rechecked with reference to the original image and the corresponding grade information, and highly uncertain cases were excluded from the valid dataset. Because the annotations were not independently produced by multiple annotators, no formal inter-annotator agreement metric was calculated. Nevertheless, the above manual quality-control procedure was adopted to reduce labeling errors and improve dataset reliability. All annotations were converted into the standard *.txt* format compatible with the YOLO framework.

Across all exposure conditions, a total of 21,536 annotated low-grade graphite ore instances were obtained. Among them, the “below” category contains 12,173 instances (56.53%), while the “low” category includes 9363 instances (43.47%), resulting in an approximate ratio of 1.3:1. To ensure reliable training and evaluation, the dataset was randomly split into training, validation, and test sets at a ratio of 8:1:1. It should be noted that the split was performed at the acquisition-batch level rather than the individual-image level. Specifically, images collected from the same conveyor sequence, acquisition session, or ore batch were assigned to only one subset and did not appear simultaneously in the training, validation, and test sets. This strategy was adopted to avoid highly similar samples from the same continuous acquisition process being distributed across different subsets, thereby reducing the risk of data leakage and overly optimistic performance evaluation. The detailed data distribution is summarized in Table 2.

Table 2. Detailed distribution of the dataset splits.

Category	Training Set	Validation Set	Test Set	Total
Total	17,198 (79.86%)	2168 (10.06%)	2170 (10.08%)	21,536 (100%)
Below	9715 (79.81%)	1233 (10.13%)	1225 (10.06%)	12,173 (56.53%)
Low	7483 (79.92%)	935 (9.99%)	945 (10.09%)	9363 (43.47%)

4.2. Experimental Environment and Hyperparameter Settings

To evaluate the effectiveness of the proposed LGG-YOLO-Seg for low-grade graphite ore instance segmentation, a unified training configuration was adopted. Specifically, the initial learning rate was set to 0.001, the batch size to 8, the number of data loading workers

to 4, and the total training epochs to 200. The model was optimized using stochastic gradient descent (SGD) to ensure stable convergence. To mitigate overfitting and reduce unnecessary training cost, an early stopping strategy was employed with a patience of 50 epochs, i.e., training was terminated if no improvement in validation performance was observed over 50 consecutive epochs. Furthermore, mosaic data augmentation [45] was disabled during the final 10 epochs to alleviate excessive data perturbation and improve convergence stability, thereby enhancing generalization performance on the test set.

All experiments were conducted on a workstation equipped with an Intel Core i5-12490F CPU (3.00 GHz; Intel Corporation, Ho Chi Minh City, Vietnam) and an NVIDIA GeForce RTX 3080 Ti GPU (NVIDIA Corporation, Santa Clara, CA, USA). The detailed hardware and software configurations are summarized in Table 3. For FPS evaluation, all models were tested under the same hardware and software environment using an input size of 640×640 pixels and a batch size of 1. The confidence threshold was set to 0.25, and the NMS IoU threshold was set to 0.7, following the default YOLO inference configuration. All inference tests were conducted using the PyTorch implementation under FP32 precision, without ONNX, TensorRT, FP16, or other additional acceleration strategies. The FPS values were calculated based on the average inference speed on the test set after warm-up runs, and the same measurement protocol was applied to all compared models to ensure fairness and reproducibility. In this study, the reported FPS includes image preprocessing, model forward inference, NMS, and mask post-processing, but does not include disk image loading time.

Table 3. Experimental Environment Configuration.

Name	Parameter
System	Windows10 (64bit)
CPU	12th-generation Intel Core i5-12490F
Memory	31 GB
GPU	NVIDIA GeForce RTX 3080 Ti
Video Memory	11 GB
Programming Software	Python 3.10.15
Deep Learning Framework	PyTorch 2.5.1
GPU Acceleration Library	CUDA11.8

4.3. Evaluation Metrics

To comprehensively and objectively validate the performance of the LGG-YOLO-Seg model in practical industrial sorting scenarios, this study established an evaluation system that considers both segmentation accuracy and network lightweight characteristics.

In terms of accuracy assessment, the core metrics include Precision (P) and the mean Average Precision at an IoU threshold of 0.50 (mAP50). Considering the particularity of instance segmentation tasks, the above accuracy metrics are further divided into four specific sub-items: P_{box} and P_{mask} , which measure the accuracy of bounding box localization and mask generation, respectively, and $mAP50_{box}$ and $mAP50_{mask}$, which reflect the overall detection and segmentation performance.

Precision mainly reflects the reliability of the network's positive predictions, i.e., the proportion of correctly predicted positive samples among all samples predicted as positive, and is calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\%. \quad (6)$$

where TP denotes the number of positive samples that are correctly predicted as low-grade graphite ores by the model, and FP represents the number of background or other objects

that are incorrectly classified as low-grade ores. In addition, mAP50 is a key metric for evaluating the overall performance of multi-class detection, which is obtained by averaging the Average Precision (AP) over all categories. The specific mathematical expression is given as follows:

$$\text{mAP}_{50} = \frac{1}{N} \sum_{c=1}^N \text{AP}_c^{0.5}. \quad (7)$$

Considering that the constructed graphite ore dataset in this study contains only two subclasses, $N = 2$. $\text{AP}_{c0.5}$ represents the average precision of the c -th category at an IoU threshold of 0.5.

In addition to accuracy, computational efficiency and the feasibility of deployment on edge devices are also key evaluation aspects. In this study, four core metrics are selected to quantify the lightweight characteristics of the network: floating-point operations (FLOPs), frames per second (FPS), the total number of model parameters (Params), and model size (Size). Among them, FPS directly reflects the inference efficiency and segmentation speed of the algorithm in processing video streams, while FLOPs focuses on evaluating the computational complexity required for a single forward inference, which is a critical factor in determining whether real-time processing can be achieved on resource-constrained edge devices [46]. The corresponding computational complexity is calculated as follows:

$$\text{GFLOPS} = W \times H \times K \times K \times \text{Cin} \times \text{Cout}. \quad (8)$$

In this definition, W and H correspond to the spatial width and height of the input feature map, respectively, while K denotes the kernel size of the applied square convolution. Cin and Cout represent the number of input channels and output channels, respectively, where the latter is also equivalent to the number of convolution kernels used in the layer.

Finally, Params and Size jointly determine the memory consumption and generalization capability of the model during operation. Effectively reducing these four lightweight metrics is a prerequisite for successfully deploying the sorting system in complex and resource-constrained industrial environments, enabling an optimal balance among hardware availability, real-time inference, and high-precision segmentation.

5. Results and Analysis

5.1. Model Training

To verify the effectiveness of the proposed LGG-YOLO-Seg model in practical applications, training was conducted on the self-constructed low-grade graphite ore dataset using the predefined hyperparameter settings, with YOLO11n-seg adopted as the baseline model for comparative analysis. During the complete training process of 200 epochs, the dynamic evolution of the mean Average Precision for segmentation masks ($\text{mAP}_{50_{\text{mask}}}$) and segmentation loss (Seg Loss) is illustrated in Figure 7.

In the initial stage of model training (approximately the first 25 epochs), both network architectures exhibit a rapid adaptation to the image features of graphite ores. During this stage, the $\text{mAP}_{50_{\text{mask}}}$ curve shows a sharp upward trend, while the corresponding segmentation loss decreases rapidly, indicating that both models possess excellent initial convergence performance.

As the training process progresses into the middle and later stages, significant differences in feature learning dynamics between the two models begin to emerge. The mask accuracy curve of the baseline YOLO11n-seg model, after reaching a high level, exhibits frequent and noticeable fluctuations, while its segmentation loss remains slightly higher than that of the improved model throughout the mid-to-late training stages. This phenomenon reflects the limitation of the baseline model in capturing features with a

single-scale receptive field when dealing with irregular shapes and complex textures of low-grade ore boundaries.

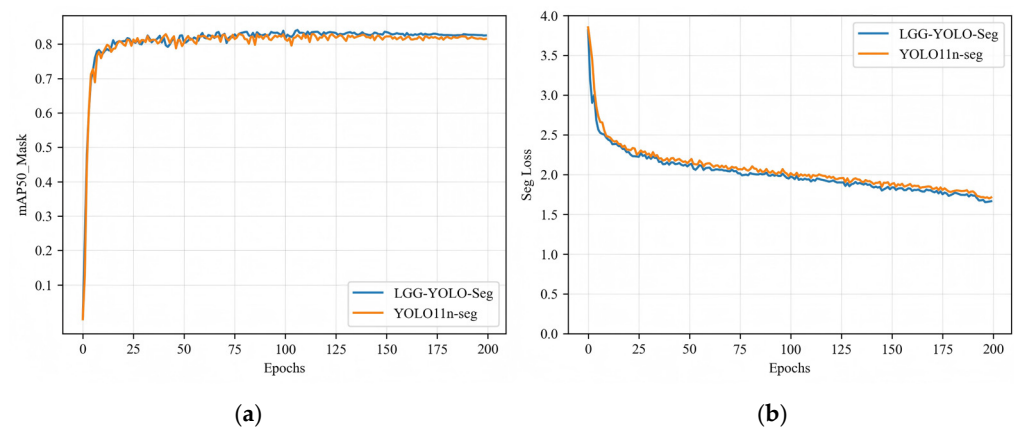


Figure 7. Training performance comparison between the LGG-YOLO-Seg model and the baseline model. (a) mAP50_{mask}-epochs curve; (b) seg loss-epochs curve.

In contrast, the proposed LGG-YOLO-Seg model demonstrates superior training robustness. As shown in the figure, the improved model not only maintains a faster and smoother performance improvement during the rising phase, but also exhibits strong convergence capability in the later stable stage. Its accuracy fluctuations are significantly suppressed, and it ultimately converges to a slightly higher accuracy level than the baseline model while maintaining the lowest overall loss. This stable and superior training performance demonstrates that the integration of a lightweight backbone and multi-scale feature extraction modules in LGG-YOLO-Seg effectively enhances the model's ability to capture deep features and adapt to low-grade graphite ores with diverse shapes in complex industrial environments.

5.2. Ablation Experiments

To systematically evaluate the contributions of each improved module in the low-grade graphite ore instance segmentation task, this study adopts the original YOLO11n-seg as the baseline model and constructs an ablation experiment framework based on the channel-scaled YOLO11cp-seg architecture, where improvement strategies are gradually introduced. All experiments are conducted under identical data splits and training parameter settings to ensure fairness and comparability of the results. On this basis, the independent effects and synergistic contributions of different components on model performance are analyzed by incrementally integrating each module, thereby comprehensively validating the effectiveness and rationality of the proposed method.

This study focuses on three key improvement directions for structural design and ablation analysis, which are summarized as follows:

- A. Introducing Linear Deformable Convolution (LDConv) into the backbone network to replace the standard convolution operations in the downsampling stages;
- B. Reconstructing the original C3k2 feature extraction module by embedding the PKI structure with multi-scale receptive field modeling capability, forming the C3k2-PKI module;
- C. For the limited-category and limited-category industrial segmentation scenario with two low-grade ore subclasses in this study, redesigning the original segmentation head with a lightweight strategy, resulting in the Segment-EfficientHead module.

The ablation experiment results are presented in Table 4. To systematically evaluate the independent contributions and synergistic effects of the proposed modules in the low-grade

graphite ore instance segmentation task, a series of ablation experiments were conducted under unified training parameters and dataset configurations.

Table 4. Results of the Ablation Experiments.

Model	A	B	C	P_{box} (%)	P_{mask} (%)	$mAP50_{box}$ (%)	$mAP50_{mask}$ (%)	FPS	FLOPS (G)	Params (M)	Size (MB)
YOLO11n-seg				90.5	85.9	92.4	83.2	82	10.4	2.84	5.83
YOLO11cp-seg				89.7	84.6	91.2	82.8	86	9.3	1.82	3.54
Transition Model	✓			91.3	86.2	92.9	83.5	85	8.3	1.48	3.17
		✓		91.0	85.6	92.6	83.1	83	8.9	1.73	3.39
			✓	90.6	85.1	92.1	83.0	88	8.2	1.45	3.05
	✓	✓		92.1	86.8	93.0	83.9	87	7.9	1.40	2.93
	✓		✓	91.8	86.5	92.7	83.7	89	7.6	1.32	2.71
		✓	✓	91.5	86.0	92.8	83.4	88	7.8	1.38	2.80
LGG-YOLO-seg	✓	✓	✓	92.6	87.4	93.3	83.6	91	7.4	1.26	2.62

When only the LDConv module (Strategy A) is introduced into the backbone network, the model significantly improves its ability to capture features of irregular boundaries and multi-scale ores while maintaining the lightweight structure. Specifically, P_{box} and P_{mask} increase to 91.3% and 86.2%, respectively, and both detection and segmentation accuracy of $mAP50$ are improved. Meanwhile, the FLOPs and model parameters are reduced to 8.3 G and 1.48 M compared with the baseline, while the inference speed remains at 85 FPS, demonstrating that LDConv effectively enhances feature representation while controlling computational complexity. Subsequently, when the C3k2-PKI module (Strategy B) is independently introduced, the network shows clear advantages in multi-scale information fusion. The $mAP50_{box}$ and $mAP50_{mask}$ values reach 92.6% and 83.1%, respectively, while maintaining reasonable inference speed and parameter scale, indicating that the expansion of multi-scale receptive fields plays an important role in recognizing complex ore structures. The results of Strategy C (integrating Segment-EfficientHead) show that although the improvement in accuracy is relatively limited, the inference speed (FPS) reaches 88, achieving significant optimization. At the same time, the model size and number of parameters are further reduced, verifying the efficiency of the lightweight segmentation head in industrial limited-category scenarios.

In the transition models with progressively integrated modules, the combination strategies A + B, A + C, and B + C all show further improvements in both accuracy and lightweight metrics. Finally, the fully integrated LGG-YOLO-Seg model (A + B + C) achieves improvements of approximately 2.1% and 1.5% in P_{box} and P_{mask} , respectively, compared with the baseline YOLO11n-seg model, while $mAP50_{box}$ and $mAP50_{mask}$ are improved by approximately 0.9% and 0.4%, respectively. At the same time, FLOPs are reduced by approximately 28.8%, the number of parameters decreases by approximately 55.6%, the model size is reduced by approximately 55.1%, and the frame rate is increased by approximately 11.0%. It should be noted that the final LGG-YOLO-Seg does not achieve the highest value in every single accuracy metric. For example, some intermediate configurations, such as A + B and A + C, obtain slightly higher $mAP50_{mask}$ values than the final model. However, these differences are very small, while the final model achieves the best overall lightweight performance, including the lowest FLOPs, the fewest parameters, the smallest model size, and the highest FPS. In addition, LGG-YOLO-Seg obtains the highest P_{box} , P_{mask} , and $mAP50_{box}$ among all ablation configurations. Therefore, the final architecture was selected not because it maximizes a single segmentation metric, but because it provides the most favorable balance among segmentation accuracy, computational cost, model

size, and inference speed. This trade-off is particularly important for real-time industrial sorting scenarios, where deployment efficiency, response speed, and resource consumption are as critical as mask accuracy. These results demonstrate that LDConv, C3k2-PKI, and Segment-EfficientHead can independently enhance different aspects of model performance, and their synergistic combination provides practical support for efficient and intelligent sorting of low-grade graphite ores in industrial scenarios.

To comprehensively evaluate the performance contributions of different modules in the low-grade graphite ore instance segmentation task, we designed strict module replacement ablation experiments and conducted fair comparisons under unified training parameters and dataset conditions. The results are shown in Table 5.

Table 5. Ablation study results for individual modules.

Model	P_{box} (%)	P_{mask} (%)	$mAP50_{box}$ (%)	$mAP50_{mask}$ (%)	FPS	FLOPS (G)	Params (M)	Size (MB)
PSConv	91.2	86.0	92.5	82.9	87	7.6	1.42	2.84
SPDConv	91.5	86.5	93.1	83.4	81	9.4	1.74	3.56
LDConv	91.3	86.2	92.9	83.5	85	8.3	1.48	3.17
C3k2-iRMB	90.7	85.3	92.2	82.8	82	8.7	1.70	3.24
C3k2-DRB	90.9	85.5	92.4	82.9	80	9.0	1.76	3.81
C3k2-PKI	91.0	85.6	92.6	83.1	83	8.9	1.73	3.39
Segment-LADH	90.4	85.0	92.0	82.8	85	8.4	1.48	3.12
Segment-LSCD	90.2	84.7	91.8	82.4	90	7.1	1.34	2.78
Segment-EfficientHead	90.6	85.1	92.1	83.0	88	8.2	1.45	3.05

Regarding the feature extraction capability of the backbone network, LDConv demonstrates superior multi-scale adaptability and irregular boundary modeling ability compared with PSConv and SPDConv. Although it is slightly inferior in lightweight metrics (with slightly higher parameters and model size), it outperforms the comparison networks in P_{box} , P_{mask} , $mAP50_{box}$ and $mAP50_{mask}$, reaching 91.3%, 86.2%, 92.9%, and 83.5%, respectively, verifying its effectiveness in capturing complex ore particle shapes and texture features. For the improvement in the C3k2 module, C3k2-PKI exhibits more sufficient feature representation capability when handling ore targets with significant scale variations. The $mAP50_{box}$ and $mAP50_{mask}$ are improved by approximately 0.2–0.4% compared with the C3k2-iRMB and C3k2-DRB modules. Meanwhile, the growth of FLOPs and parameters is effectively suppressed through depthwise separable convolution and pointwise fusion strategies, achieving a balance between accuracy and lightweight design. In terms of the segmentation head, Segment-EfficientHead is structurally optimized for limited-category industrial scenarios. While maintaining segmentation accuracy, it significantly improves inference speed and reduces the number of parameters and model size. Compared with Segment-LADH and Segment-LSCD, it is more efficient, fully demonstrating the advantages of lightweight design.

Overall, LDConv enhances backbone feature representation and multi-scale adaptability, C3k2-PKI strengthens feature transmission and global information capture, and Segment-EfficientHead significantly optimizes the computational cost and inference speed of the segmentation head. Although the final LGG-YOLO-Seg does not maximize every individual accuracy indicator, it achieves the most practical overall trade-off among accuracy, computational complexity, model size, and real-time performance. Therefore, the final architecture is preferable for industrial deployment, where fast response, low resource consumption, and stable segmentation performance are all required for intelligent sorting of low-grade graphite ores.

5.3. Comparison Study

To validate the overall performance of the proposed LGG-YOLO-Seg model in the low-grade graphite ore instance segmentation task, we conducted systematic comparative experiments with multiple mainstream segmentation models. To ensure a controlled and reproducible comparison, all models were evaluated under the same dataset split, input size, training epochs, hardware environment, and evaluation metrics. No pre-trained weights were used, and the main training settings were kept consistent across models. This protocol was adopted to reduce confounding factors and evaluate the relative performance of different models under the same experimental budget. It should be noted that this setting does not necessarily represent the individually optimized performance of each architecture, since different models may benefit from architecture-specific hyperparameter tuning. Therefore, the comparison results should be interpreted as performance under a unified experimental protocol rather than the absolute upper-bound performance of each baseline model.

The comprehensive comparison results are presented in Table 6, and Figure 8 illustrates the comparison of different models in terms of lightweight performance metrics. To further clarify the advantages of the proposed method, the comparison includes both high-accuracy segmentation frameworks and lightweight YOLO-based segmentation models, thereby evaluating the trade-off among accuracy, computational cost, model size, and inference speed.

Table 6. Comparative experimental results.

Model	P_{box} (%)	P_{mask} (%)	$mAP50_{box}$ (%)	$mAP50_{mask}$ (%)	FPS	FLOPS (G)	Params (M)	Size (MB)
Mask R-CNN [29]	92.8	87.5	93.5	83.9	45	176.3	43.7	214
U-Net [28]	93.1	88.3	93.8	84.5	62	92.8	30.2	96
YOLOv8n-seg [47]	90.8	85.7	92.8	83.2	78	12.0	3.27	6.51
YOLOv9c-seg [48]	92.7	87.2	93.6	83.7	53	156.8	27.5	53.4
YOLOv12n-seg [49]	90.2	84.8	92.1	82.4	81	10.4	2.82	5.79
YOLOv13n-seg [50]	90.7	85.4	92.6	83.0	82	10.3	2.73	5.74
YOLO11s-seg	92.1	86.1	92.9	83.4	71	35.6	10.08	19.8
YOLO11n-seg	90.5	85.9	92.4	83.2	82	10.4	2.84	5.83
YOLO11cp-seg	89.7	84.6	91.2	82.8	86	9.3	1.82	3.54
LGG-YOLO-seg	92.6	87.4	93.3	83.6	91	7.4	1.26	2.62

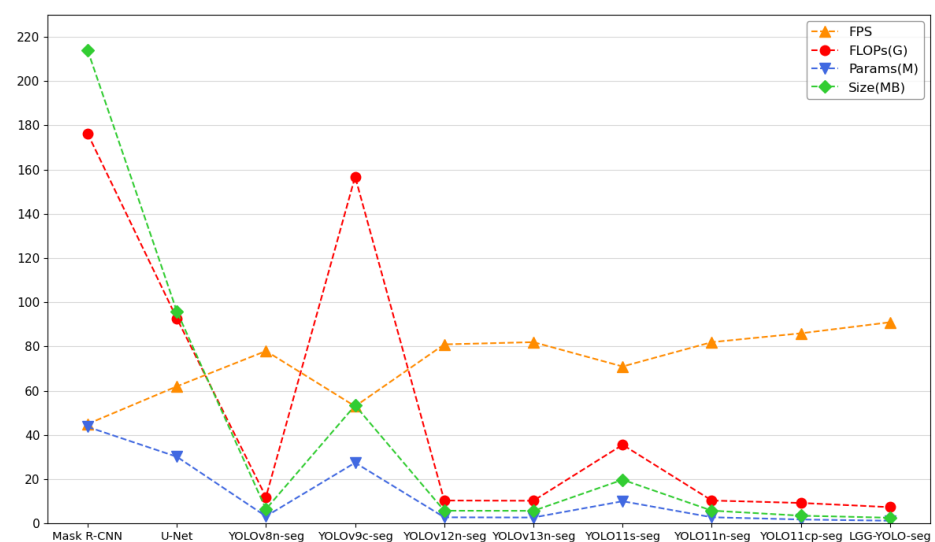


Figure 8. Comparison of different models in terms of lightweight performance metrics.

Traditional two-stage segmentation frameworks, such as Mask R-CNN and the classic U-Net, demonstrate high segmentation accuracy, with Mask R-CNN achieving $mAP50_{box}$ and $mAP50_{mask}$ of 93.5% and 83.9%, respectively, and U-Net achieving 93.8% and 84.5%. However, these models exhibit very high computational complexity and inference speeds of only 45–62 FPS, limiting their feasibility for deployment in resource-constrained or real-time industrial scenarios. Although these models can provide competitive accuracy, their large parameter sizes, high FLOPs, and relatively slow inference speed make them less suitable for on-site low-grade graphite ore sorting, where both real-time response and edge deployment are required.

In contrast, lightweight YOLO-based segmentation models show clear advantages in inference speed and model compression. For example, YOLOv8n-seg achieves an FPS of 78, with 3.27 M parameters and a model size of 6.51 MB, though its $mAP50_{box}$ and $mAP50_{mask}$ are slightly lower at 92.8% and 83.2%. YOLOv12n-seg and YOLOv13n-seg also maintain relatively compact structures, but their segmentation accuracy remains lower than that of LGG-YOLO-Seg on the low-grade graphite ore dataset. This indicates that general lightweight segmentation models may not fully address the challenges of irregular ore boundaries, multi-scale particle distributions, and complex surface textures.

Further analysis of the YOLO11 series shows that LGG-YOLO-Seg achieves significant overall performance gains: P_{box} and P_{mask} reach 92.6% and 87.4%, while $mAP50_{box}$ and $mAP50_{mask}$ reach 93.3% and 83.6%. It maintains high accuracy while being highly lightweight, with FLOPs of only 7.4 G, 1.26 M parameters, a model size of 2.62 MB, and high-speed inference at 91 FPS. Compared to the original YOLO11n-seg, LGG-YOLO-Seg shows improvements across all three core accuracy metrics while reducing computational cost by approximately 28.8%, parameters by 55.6%, and model size by 55.1%, and increasing FPS by approximately 11.0%, achieving an optimal balance among accuracy, lightweight design, and inference speed.

In addition, compared with the lightweight YOLOv11cp-seg model of the same category, LGG-YOLO-Seg also demonstrates clear advantages. This improvement indicates that the proposed method is not a simple model scaling or compression strategy. Instead, it introduces task-oriented structural optimization for low-grade graphite ore segmentation. Its improved backbone with LDConv enhances the capability of capturing features of irregular ore structures, the C3k2-PKI module strengthens multi-scale information representation, and the Segment-EfficientHead significantly optimizes inference speed and computational cost while maintaining segmentation accuracy. Compared with similar lightweight segmentation methods, LGG-YOLO-Seg achieves a better balance between boundary-sensitive segmentation accuracy and deployment efficiency, which is particularly important for real-time intelligent sorting of low-grade graphite ores.

These results fully indicate that LGG-YOLO-Seg achieves substantial model compression and acceleration while ensuring high accuracy, providing a feasible technical solution and lightweight scheme for real-time intelligent sorting of low-grade graphite ores in industrial scenarios.

5.4. Visualization Experiments

To visually demonstrate the performance of the LGG-YOLO-Seg model in the low-grade graphite ore instance segmentation task, this study selects three representative test images under exposure conditions of 400 μs , 600 μs , and 800 μs , respectively, to analyze the visualization results. Figure 9 presents a comparison between LGG-YOLO-Seg and the baseline YOLO11n-seg model. Low-grade graphite ores are divided into two subcategories and annotated with different colors: below-average grade (blue) and low grade (green). The confidence score (0–1) within each bounding box reflects the reliability of each prediction.

Red arrows are used to highlight key regions, facilitating the comparison of segmentation performance between different models.

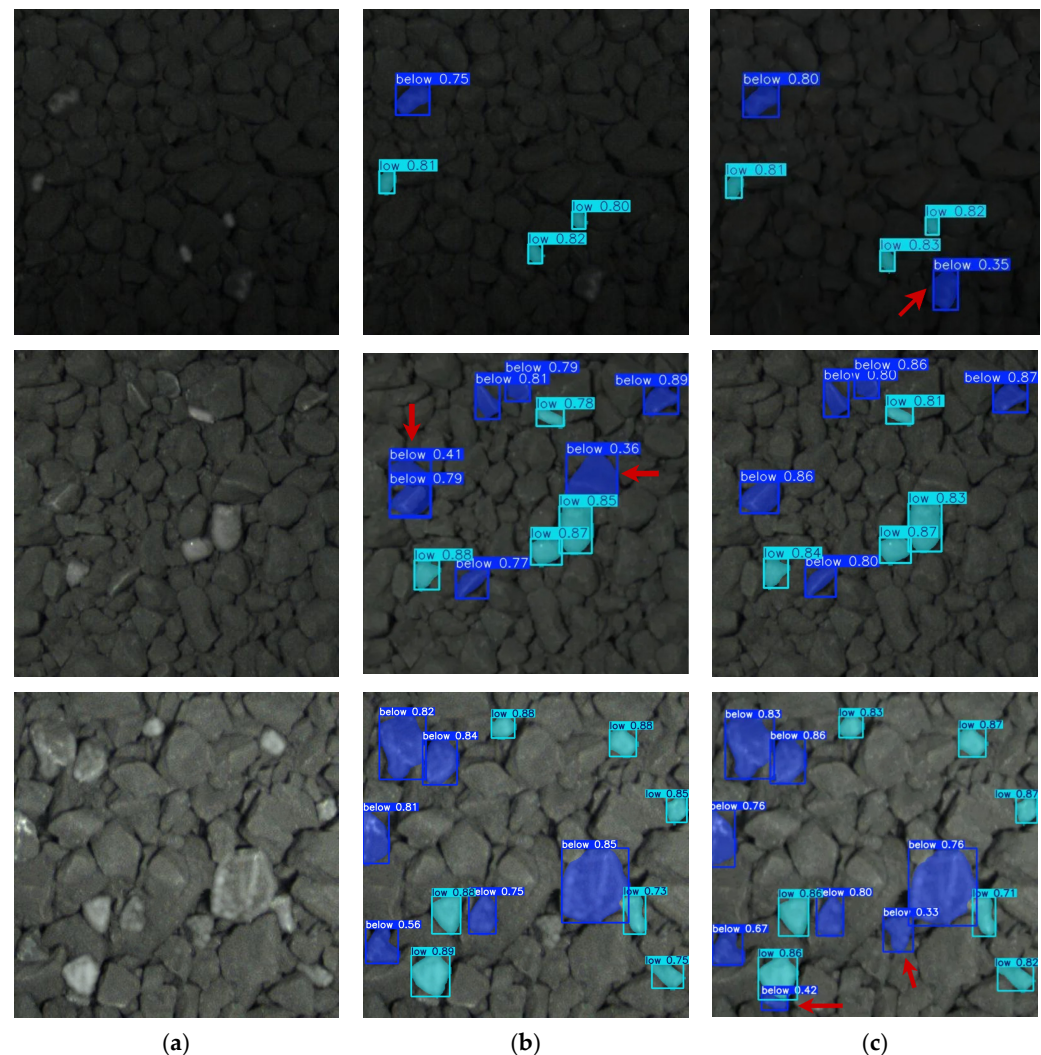


Figure 9. Visual comparison of model segmentation results. Red arrows indicate key comparison regions. (a) Original Image; (b) YOLO11n-seg; (c) LGG-YOLO-Seg.

In the first row (400 μ s exposure), the visual features of low-grade ore blocks are relatively clear. However, the baseline model exhibits missed detections, while LGG-YOLO-Seg accurately identifies and fully segments these ore blocks, achieving better confidence scores and segmentation accuracy, demonstrating its robustness under clear conditions. In the second row (600 μ s exposure), low-grade ores appear densely stacked. YOLO11n-seg misclassifies some high-grade ores, whereas LGG-YOLO-Seg correctly identifies and segments these targets, showing its capability in handling complex ore arrangements. In the third row (800 μ s exposure), complex illumination interference is introduced. The baseline model can only detect relatively large low-grade ore blocks, while LGG-YOLO-Seg still achieves accurate segmentation under the same conditions, verifying its robustness and adaptability under varying lighting and interference conditions.

The visualization results show that LGG-YOLO-Seg exhibits stable and reliable segmentation performance under different exposure conditions, with overall performance significantly superior to the baseline YOLO11n-seg model. The model maintains high recognition accuracy and segmentation integrity across various scenarios, including clear

conditions, densely distributed targets, and complex illumination interference, effectively reducing missed detections and misclassifications.

This performance improvement is mainly attributed to the enhanced representation capability of the improved backbone LDConv for complex texture features, as well as the optimized multi-scale information interaction provided by the C3k2-PKI module. Meanwhile, the Segment-EfficientHead improves boundary detail representation while reducing computational cost, enabling the model to maintain strong segmentation accuracy under lightweight constraints.

Overall, LGG-YOLO-Seg demonstrates excellent robustness and stability in complex industrial environments, maintaining reliable segmentation performance under various interference conditions and showing strong potential for practical engineering applications.

Although the proposed model demonstrates stable segmentation performance in most representative scenarios, several failure cases were still observed under more challenging conditions. For example, when ore particles are severely occluded or densely adhered, the visible boundaries between adjacent instances become incomplete, which may lead to inaccurate mask separation or missed detection of small ore fragments. In addition, under extreme illumination conditions, such as strong shadows or local overexposure, the surface texture and grayscale contrast of graphite ores are weakened, making it difficult for the model to distinguish ore boundaries from the background. Misclassification may also occur when different low-grade ore categories exhibit highly similar surface textures and colors, especially under dust coverage or uneven lighting. These failure cases indicate that although LDConv and C3k2-PKI improve boundary adaptation and multi-scale feature extraction, the current model still mainly relies on visible surface features and may be limited when visual cues are seriously degraded. Future work will further enrich the dataset with more difficult samples and introduce multimodal information to improve robustness under severe occlusion, extreme lighting, and similar-texture categories.

5.5. Generalization Experiments

To further evaluate the adaptability of the proposed LGG-YOLO-Seg model across different mineral image scenarios, a cross-dataset validation experiment was conducted on the publicly available Actinolite Dataset [51]. This dataset contains 1092 microscopic images with a total of 6244 annotated instances, covering multiple mineral types including actinolite, garnet, and hornblende. During the experiments, all models followed the same training strategies and evaluation criteria as used on the graphite ore dataset, ensuring objective and consistent comparative results. The quantitative results are summarized in Table 7, and the corresponding visual comparisons are shown in Figure 10, with key regions indicated by red arrows to clearly highlight differences in segmentation performance among models for complex mineral structures.

Table 7. Comparative results on the Actinolite Dataset.

Model	mAP50 _{box} (%)	mAP50 _{mask} (%)	FPS	FLOPS (G)	Params (M)	Size (MB)
YOLO11n-seg	89.7	66.5	83	10.4	2.84	5.85
YOLO11cp-seg	89.3	66.2	85	9.3	1.82	3.53
LGG-YOLO-seg	90.8	67.3	89	7.4	1.26	2.63

The experimental results indicate that LGG-YOLO-Seg maintains stable performance on this dataset. Compared with the baseline YOLO11n-seg model, it achieves improvements of approximately 1.4% in mAP50_{box} and 1.2% in mAP50_{mask} while simultaneously reducing model complexity: parameters decreased by about 52.1%, FLOPs reduced by approximately 18.6%, model size shrank by about 49.7%, and inference speed increased by around 8.8%.

These results suggest that the proposed LGG-YOLO-Seg model has certain cross-dataset adaptability and can maintain a favorable efficiency–accuracy trade-off on the evaluated Actinolite Dataset. LGG-YOLO-Seg also shows promising adaptability and stability in the evaluated cross-dataset scenario, indicating its potential for broader industrial inspection applications, although further validation on more datasets is still needed.

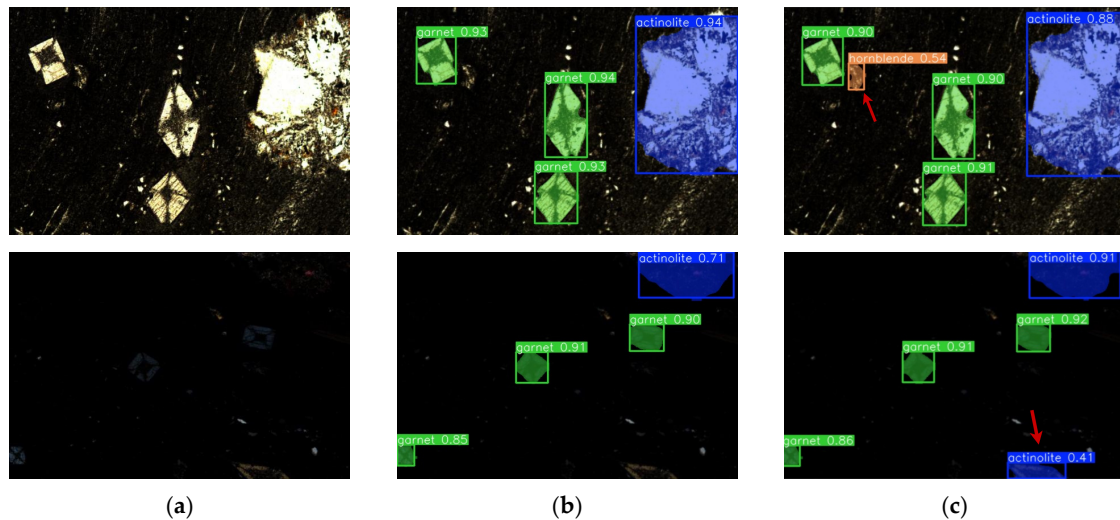


Figure 10. Visual comparison on the Actinolite Dataset. (a) Original Image; (b) YOLO11n-seg; (c) LGG-YOLO-Seg.

5.6. Limitations

Compared with existing instance segmentation methods, although LGG-YOLO-Seg achieves a good balance among accuracy, lightweight design, and inference efficiency in the low-grade graphite ore-sorting task, it still has certain limitations that require further study and improvement:

- Although the dataset constructed in this study incorporates multi-exposure acquisition and fine-grained annotation to simulate illumination variations in real industrial environments, its overall scale and scene diversity remain limited. In particular, the images were mainly collected from a single mining area, namely China Minmetals Group (Heilongjiang) Graphite Industry Co., Ltd., under relatively specific acquisition conditions. Therefore, the generalization ability of LGG-YOLO-Seg across different mining areas, ore types, production backgrounds, and more complex illumination disturbances, such as shadows, dust, uneven lighting, and surface contamination, still requires further validation.
- The current method mainly relies on visible-light image information for ore-grade classification. Although visible-light images can effectively capture surface texture, color, and boundary features of graphite ores, they may be insufficient for distinguishing samples with similar appearances or subtle differences in internal mineral composition. This limitation may affect the model's discrimination capability when ore surfaces are affected by dust, adhesion, shadow occlusion, or complex industrial lighting. In future work, multimodal data such as hyperspectral, near-infrared, or X-ray imaging can be incorporated to enhance the representation of both surface and internal ore features.
- Although LGG-YOLO-Seg significantly reduces parameter size and computational complexity through structural compression and module optimization, and demonstrates high inference efficiency in theoretical metrics (e.g., FLOPs, parameters, and model size) and experimental environments, systematic deployment and long-term evaluation on real industrial edge devices (such as embedded GPUs or on-site sort-

ing systems) remain insufficient. In addition, the actual memory occupation and power consumption of the current model on typical edge hardware have not yet been quantitatively evaluated. Therefore, although the lightweight indicators suggest good deployment potential, its real-time performance, resource consumption, power efficiency, operational stability, and environmental adaptability in different production scenarios still require further validation and improvement to better support large-scale industrial applications.

Future work will focus on expanding the scale and diversity of the dataset by collecting ore images from different mining areas, ore types, production lines, and environmental conditions, especially under challenging scenarios involving shadows, dust, uneven illumination, and background interference. In addition, multimodal information will be further integrated, and deployment validation on real industrial devices will be strengthened. During this process, quantitative evaluation of peak memory usage, power consumption, inference latency, and long-term operating stability on representative edge hardware will also be conducted. These efforts are expected to improve model stability, generalization capability, environmental adaptability, and practical engineering value in complex mineral-sorting scenarios.

6. Conclusions

Efficient identification and sorting of low-grade graphite ores are of great significance for improving resource utilization, ensuring product quality, and reducing energy consumption and environmental impact during production. However, traditional sorting methods in practice often rely on complex equipment or manual experience, which are characterized by low efficiency, high cost, and limited adaptability, making them insufficient to meet the demands of intelligent and green mining. To address these issues, this study proposes a lightweight instance segmentation method, LGG-YOLO-Seg, for rapid identification and intelligent sorting of low-grade graphite ores in complex industrial environments.

At the data level, this study constructs a graphite ore image dataset based on real industrial scenarios. The dataset is collected under multi-exposure conditions and finely annotated, covering two grades of low-grade ore samples, providing a reliable foundation for model training and evaluation. At the methodological level, the network structure is systematically optimized to build a more compact and efficient backbone. Linear deformable convolution is introduced into the backbone to enhance feature modeling capability for irregularly shaped ores, the feature extraction module is reconstructed to improve multi-scale information fusion, and a lightweight segmentation head is designed to reduce computational overhead and enhance inference efficiency.

Meanwhile, the effectiveness of each proposed module is validated through ablation and comparative experiments. The experimental results demonstrate that, compared with the baseline YOLO11n-seg model, the proposed LGG-YOLO-Seg achieves comparable and slightly improved segmentation accuracy while substantially reducing model complexity and resource consumption. Specifically, the main advantage of LGG-YOLO-Seg lies not in a large absolute improvement in segmentation accuracy, but in achieving a more favorable efficiency–accuracy trade-off, with reduced FLOPs, fewer parameters, smaller model size, and improved inference speed. This indicates that the proposed method is more suitable for real-time industrial sorting and resource-constrained edge deployment scenarios.

In summary, LGG-YOLO-Seg demonstrates competitive segmentation performance and strong lightweight deployment potential in the low-grade graphite ore instance segmentation task. Future work will focus on dataset expansion, multimodal information fusion, and optimization for deployment on industrial devices, further improving model stability and practicality in complex environments. More specifically, future research will

further explore parameter-level architectural optimization, hardware-aware acceleration, and closed-loop validation in real industrial sorting systems, thereby extending the applicability of the proposed method to more complex and diverse mineral-sorting scenarios. These efforts are expected to provide technical support for intelligent and sustainable mining development.

Author Contributions: Conceptualization, J.W., Y.Y. and Z.S.; methodology, J.W.; software, J.W. and G.Y.; validation, Z.S. and Z.Q.; resources, J.W. and Y.Y.; data curation, J.W. and Z.Q.; writing—original draft preparation, J.W. and Z.S.; writing—review and editing, J.W.; visualization, G.Y.; supervision, Z.S.; project administration, Z.S.; funding acquisition, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National key Research and Development program of China 2020YFB1713700.

Institutional Review Board Statement: Not applicable. This study did not involve humans or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data supporting the results of this study are available from the corresponding author upon request.

Acknowledgments: The authors would like to express their sincere gratitude to China Minmetals Corporation Graphite Industry Co., Ltd., for generously providing the site access and mineral raw materials that supported this research.

Conflicts of Interest: Author Jionghui Wang, Yuxing Yu and Guangzheng Yao were employed by the company China Minmetals Corporation (Heilongjiang) Graphite Industry Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Zhang, J.; Liang, C.; Dunn, J.B. Graphite flows in the US: Insights into a key ingredient of energy transition. *Environ. Sci. Technol.* **2023**, *57*, 3402–3414. [[CrossRef](#)] [[PubMed](#)]
2. Keeling, J.L. Graphite: Properties, uses and South Australian resources. *MESA J.* **2017**, *84*, 28–41.
3. Lin, K.; Fang, H.; Gao, A.; Yu, H.; Wang, L. Nanoburl graphites. *Adv. Mater.* **2021**, *33*, 2007513. [[CrossRef](#)]
4. Andersen, S.Z.; Čolić, V.; Yang, S.; Schwalbe, J.A.; Nielander, A.C.; McEnaney, J.M.; Enemark-Rasmussen, K.; Baker, J.G.; Singh, A.R.; Rohr, B.A.; et al. A rigorous electrochemical ammonia synthesis protocol with quantitative isotope measurements. *Nature* **2019**, *570*, 504–508. [[CrossRef](#)] [[PubMed](#)]
5. Rezapour, M.R.; Myung, C.W.; Yun, J.; Ghassami, A.; Li, N.; Yu, S.U.; Hajibabaei, A.; Park, Y.; Kim, K.S. Graphene and graphene analogs toward optical, electronic, spintronic, green-chemical, energy-material, sensing, and medical applications. *ACS Appl. Mater. Interfaces* **2017**, *9*, 24393–24406. [[CrossRef](#)] [[PubMed](#)]
6. Yadav, R.; Sharma, A.K.; Sharma, S. Advance Development in Natural Graphite Material and Its Applications: A Review. *Min. Metall. Explor.* **2025**, *42*, 361–385.
7. Wang, H.; Liu, Z.; Qu, F.; Wang, L.; Yue, X.; Zhang, X.; Shao, A. Development of Online Detection Technologies for Ore Grade. *Strateg. Study Chin. Acad. Eng.* **2024**, *26*, 152–163. [[CrossRef](#)]
8. Bulayani, M.M.; Raghupatruni, P.; Mamvura, T.; Danha, G. Exploring low-grade iron ore beneficiation techniques: A comprehensive review. *Minerals* **2024**, *14*, 796. [[CrossRef](#)]
9. Gairola, S.U.; Khanduri, A.K.; Bhuvaneshwari, V. Sustainable mining: Reducing waste and enhancing resource efficiency. *Discov. Civ. Eng.* **2025**, *2*, 75. [[CrossRef](#)]
10. Balasubramanian, A. Overview of mineral processing methods. In *Centre for Advanced Studies in Earth Science*; University of Mysore: Mysore, India, 2015.
11. Ruan, Y.; He, D.; Chi, R. Review on beneficiation techniques and reagents used for phosphate ores. *Minerals* **2019**, *9*, 253. [[CrossRef](#)]
12. Tsangaratos, P.; Ilia, I.; Spanoudakis, N.; Karageorgiou, G.; Perraki, M. Machine learning approaches for real-time mineral classification and educational applications. *Appl. Sci.* **2025**, *15*, 1871. [[CrossRef](#)]

13. Li, Y.; Yan, H.; Li, D.; Wang, H. Robust Miner Detection in Challenging Underground Environments: An Improved YOLOv11 Approach. *Appl. Sci.* **2024**, *14*, 11700. [\[CrossRef\]](#)
14. Shrivastava, A. Unmanned aerial vehicles (UAV) in mining sector: Enhancing productivity and safety. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **2025**, *239*, 2219–2235. [\[CrossRef\]](#)
15. Zhang, D.; Qian, X.; Shi, C.; Zhang, Y.; Qian, Y.; Zhou, S. Iron Ore Image Recognition Through Multi-View Evolutionary Deep Fusion Method. *Future Internet* **2025**, *17*, 553. [\[CrossRef\]](#)
16. Azhari, F.; Sennersten, C.C.; Lindley, C.A.; Sellers, E. Deep learning implementations in mining applications: A compact critical review. *Artif. Intell. Rev.* **2023**, *56*, 14367–14402. [\[CrossRef\]](#)
17. Peukert, D.; Xu, C.; Dowd, P. A review of sensor-based sorting in mineral processing: The potential benefits of sensor fusion. *Minerals* **2022**, *12*, 1364. [\[CrossRef\]](#)
18. Monecke, T.; Köhler, S.; Kleeberg, R.; Herzig, P.M.; Gemmell, J.B. Quantitative phase-analysis by the Rietveld method using X-ray powder-diffraction data: Application to the study of alteration halos associated with volcanic-rock-hosted massive sulfide deposits. *Can. Mineral.* **2001**, *39*, 1617–1633. [\[CrossRef\]](#)
19. Ribeiro, T.M.G.; Brandão, P.R.G. Development and validation of graphitic carbon analysis of graphite ore samples. *Tecnol. Em Metal. Mater. Min.* **2018**, *14*, 183–189. [\[CrossRef\]](#)
20. Cui, B.; Pan, W.; Yu, Y.; Chen, H.; Tang, Y.; Liao, M.; Wei, K.; Pan, S.; Fu, L. Experimental Study on the Washability of a Flake Graphite Ore. *J. Non-Met. Miner. Ind. Des. Res. Inst.* **2025**, *4*, 74–77.
21. Sader, J.A.; Gravel, J.; Janke, L.; Hall, L. In-depth study on carbon speciation focussed on graphite. In Proceedings of the Symposium on Critical and Strategic Materials, British Columbia Geological Survey Paper, Victoria, Columbia, 13–14 November 2015; Volume 3, p. 187.
22. Luo, X.; He, K.; Zhang, Y.; He, P.; Zhang, Y. A review of intelligent ore sorting technology and equipment development. *Int. J. Miner. Metall. Mater.* **2022**, *29*, 1647–1655. [\[CrossRef\]](#)
23. Shatwell, D.G.; Murray, V.; Barton, A. Real-time ore sorting using color and texture analysis. *Int. J. Min. Sci. Technol.* **2023**, *33*, 659–674. [\[CrossRef\]](#)
24. Liu, Y.; Zhang, Z.; Liu, X.; Wang, L.; Xia, X. Ore image classification based on small deep learning model: Evaluation and optimization of model depth, model structure and data size. *Miner. Eng.* **2021**, *172*, 107020. [\[CrossRef\]](#)
25. Liu, Y.; Wang, X.; Zhang, Z.; Deng, F. A review of deep learning in image classification for mineral exploration. *Miner. Eng.* **2023**, *204*, 108433. [\[CrossRef\]](#)
26. Ferreira, B.A.P.; Augusto, K.S.; Iglesias, J.C.A.; Caldas, T.D.P.; Santos, R.B.M.; Paciornik, S. Instance segmentation of quartz in iron ore optical microscopy images by deep learning. *Miner. Eng.* **2024**, *211*, 108681. [\[CrossRef\]](#)
27. Tang, H.; He, L.; Feng, Q.; Huang, B.; Huang, S.; Li, D. Fine-grained mineral recognition and segmentation in metal mineral optical images based on improved YOLOv8n model. *Miner. Eng.* **2025**, *227*, 109290. [\[CrossRef\]](#)
28. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241. [\[CrossRef\]](#)
29. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*; IEEE: Piscataway, NJ, USA, 2017; pp. 2961–2969. [\[CrossRef\]](#)
30. Bolya, D.; Zhou, C.; Xiao, F.; Lee, Y.J. Yolact: Real-time instance segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; IEEE: Piscataway, NJ, USA, 2019; pp. 9157–9166. [\[CrossRef\]](#)
31. Xiang, J.; Shi, H.; Huang, X.; Chen, D. Improving graphite ore grade identification with a novel FRCNN-PGR method based on deep learning. *Appl. Sci.* **2023**, *13*, 5179. [\[CrossRef\]](#)
32. Qiu, Z.; Huang, X.; Li, S.; Wang, J. Stellar-YOLO: A Graphite Ore Grade Detection Method Based on Improved YOLO11. *Symmetry* **2025**, *17*, 966. [\[CrossRef\]](#)
33. Wang, C.Y.; Liao, H.Y.M. YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems. *arXiv* **2024**, arXiv:2408.09332. [\[CrossRef\]](#)
34. Khanam, R.; Hussain, M. Yolov11: An overview of the key architectural enhancements. *arXiv* **2024**, arXiv:2410.17725. [\[CrossRef\]](#)
35. Ultralytics. YOLO11. 2024. Available online: <https://docs.ultralytics.com/zh/models/yolo11/> (accessed on 27 December 2025).
36. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2019; pp. 6105–6114. [\[CrossRef\]](#)
37. Gong, X.; Yu, J.; Zhang, H.; Dong, X. Aed-yolo11: A small object detection model based on yolo11. *Digit. Signal Process.* **2025**, *166*, 105411. [\[CrossRef\]](#)
38. Mumuni, A.; Mumuni, F. CNN architectures for geometric transformation-invariant feature representation in computer vision: A review. *SN Comput. Sci.* **2021**, *2*, 340. [\[CrossRef\]](#)
39. Zhang, X.; Song, Y.; Song, T.; Yang, D.; Ye, Y.; Zhou, J.; Zhang, L. LDConv: Linear deformable convolution for improving convolutional neural networks. *Image Vis. Comput.* **2024**, *149*, 105190. [\[CrossRef\]](#)

40. Cao, X.; Hu, Y.; Zhang, H. LKD-YOLOv8: A Lightweight Knowledge Distillation-Based Method for Infrared Object Detection. *Sensors* **2025**, *25*, 4054. [[CrossRef](#)] [[PubMed](#)]
41. Xiao, R.; Wang, H.; Wang, L.; Yuan, H. C3Ghost and C3k2: Performance study of feature extraction module for small target detection in YOLOv11 remote sensing images. In *Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024)*; SPIE: Bellingham, WA, USA, 2025; Volume 13550, pp. 464–470. [[CrossRef](#)]
42. Chen, X.; Qian, W.; Jiang, N. Improving C3K2 with RepConvBlock for high-accuracy and fast pig behavior object detection. In *International Conference on Computational Intelligence and Image Analysis (ICCIIA 2025)*; SPIE: Bellingham, WA, USA, 2025; Volume 13700, pp. 91–96. [[CrossRef](#)]
43. Chen, L.C.; Hermans, A.; Papandreou, G.; Schroff, F.; Wang, P.; Adam, H. Masklab: Instance segmentation by refining object detection with semantic and direction features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: Piscataway, NJ, USA, 2018; pp. 4013–4022. [[CrossRef](#)]
44. Human Signal. LabelImg. Available online: <https://github.com/HumanSignal/labelImg> (accessed on 20 March 2025).
45. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934. [[CrossRef](#)]
46. Liu, R.; Zhang, X.; Jin, S.; Wang, Q.; Zeng, L.; Liao, J. A Small Target Detection Model Based on an Improved RT-DETR. In *2024 4th International Conference on Industrial Automation, Robotics and Control Engineering (IARCE)*; IEEE: Piscataway, NJ, USA, 2024; pp. 434–438. [[CrossRef](#)]
47. Jocher, G. YOLOv8. 2023. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 12 January 2026).
48. Wang, C.Y.; Yeh, I.H.; Mark Liao, H.Y. Yolov9: Learning what you want to learn using programmable gradient information. In *European Conference on Computer Vision*; Springer Nature: Cham, Switzerland, 2024; pp. 1–21. [[CrossRef](#)]
49. Tian, Y.; Ye, Q.; Doermann, D. Yolov12: Attention-centric real-time object detectors. *arXiv* **2025**, arXiv:2502.12524. [[CrossRef](#)]
50. Lei, M.; Li, S.; Wu, Y.; Hu, H.; Zhou, Y.; Zheng, X.; Ding, G.; Du, S.; Wu, Z.; Gao, Y. Yolov13: Real-time object detection with hypergraph-enhanced adaptive visual perception. *arXiv* **2025**, arXiv:2506.17733. [[CrossRef](#)]
51. Actinolite Dataset. Available online: <https://github.com/xiaozesaigao/Actinolite-Thin-Section-Microscopic-Image-Dataset> (accessed on 18 February 2026).

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.