

基于 YOLOv8 和 SE 注意力机制的智能识别矿物质检测系统

时小芳 杨志远* 王硕旋

广东理工学院, 中国·广东 肇庆 526100

摘要: 针对传统矿物识别方法存在的准确率和效率问题, 本研究提出一种基于 YOLOv8 的深度学习模型, 并结合注意力机制以提升矿物自动识别的准确性和速度。该模型通过引入通道注意力和空间注意力模块, 有效增强了特征提取能力, 使得网络能够更加关注于矿物的关键特征信息, 从而提高了识别精度。实验结果显示, 与传统 YOLO 系列模型相比, 加入注意力机制的 YOLOv8 在多个标准矿物数据集上均表现出更高的 mAP (平均精确度均值) 和更低的推理时间, 证明了其在矿物识别任务上的有效性和先进性。本研究不仅为矿物快速准确识别提供了新的技术方案, 也为其他领域的目标检测任务提供了参考, 展示了深度学习和注意力机制结合的强大潜力。

关键词: 矿物识别; YOLOv8; SE 注意力机制; 深度学习; 目标检测

Intelligent recognition mineral detection system based on YOLOv8 and SE attention mechanism

Shi Xiaofang, Yang Zhiyuan*, Wang Shuoxuan

Guangdong University of Technology, China Guangdong Zhaoqing 526100

Abstract: Aiming at the accuracy and efficiency issues of traditional mineral identification methods, this study proposes a deep learning model based on YOLOv8 and combines an attention mechanism to enhance the accuracy and speed of automatic mineral identification. By introducing channel attention and spatial attention modules, the model effectively enhances the feature extraction capability, enabling the network to focus more on the key feature information of minerals and thereby improving the recognition accuracy. results show that, compared with traditional YOLO series models, the YOLOv8 with an attention mechanism achieves higher mAP (mean average precision) and lower inference time on multiple standard mineral datasets, demonstrating its effectiveness and advancement in mineral identification tasks. This research not only provides a new technical solution for rapid and accurate mineral identification but also offers a reference for target detection tasks in other fields, showcasing the powerful potential of combining deep learning and attention mechanisms.

Keywords: Mineral identification; YOLOv8; SE attention mechanism; Deep learning; Object detection

0 引言

矿产资源是国家经济发展与产业安全的重要物质基础, 然而, 随着全球矿业竞争加剧、环保要求趋严以及关键矿产资源战略地位不断提升, 我国矿产资源供给安全面临严峻挑战。面对矿产资源勘探周期长、成本高、传统检测方法效率低、依赖人工、准确性不足等问题, 发展智能化的矿物检测技术成为提升资源勘探开发效率的迫切需求。本文基于 YOLOv8 深度学习模型, 引入 SE 注意力机制, 构建一套智能矿物识别系统, 旨在实现矿物的快速与精准识别, 为矿产资源勘查与开发提供关键技术支撑, 对保障

国家资源安全、推动矿业高质量发展具有重要意义。

1 研究背景与挑战

矿物资源作为现代工业的基石, 其分选精度直接影响矿产利用效率和产业价值。传统人工分选依赖目视鉴别, 平均准确率仅 70% - 85%, 面对纹理相近矿物误判率超 35%, 不仅造成资源浪费, 还伴随高额人力成本与安全隐患。现有机器视觉技术也难以满足工业需求: 实验室检测耗时长, 自动化设备成本高昂, 传统算法对细微特征差异区分能力弱, 在光照波动、粉尘干扰等复杂环境下模型性能下降显著。深度学习虽在图像识别中表现优异, 但在矿

物分选应用中仍存在明显瓶颈：细粒度识别能力不足，难以捕捉解理纹、光泽等微小差异；高精度模型算力要求高，轻量化模型又易遗漏稀有矿物；动态环境下适配性差，高速产线中精度损失严重。

针对以上问题，本研究提出“轻量化架构 + 通道注意力增强”联合方案。通过集成双池化 SE 模块增强关键特征响应，结合 BiFPN 优化多尺度融合，在 30 类矿物识别中实现 95.2% 的准确率，较 YOLOv8 提升 6.5%，稀有矿物检测 AP 提升 18.2%。模型经剪枝与量化后体积仅 18MB，边缘推理速度达 50ms/ 样本，效率提升 40 - 60 倍，且在粉尘、低光照等恶劣环境下误差控制在 5% 以内，支持 5 米 / 秒产线速度下的 60FPS 实时分选，为矿产分选提供高精度、高效率、高鲁棒性的工业级解决方案。

2 系统设计与方法

2.1 系统架构

本系统面向矿物识别需求，设计并实现了一套端到端的轻量化智能检测平台。整体采用经典的前 - 后端解耦架构：前端基于 Vue 3 与 Element-Plus 构建单页应用，负责用户交互、图像预览与结果可视化；后端以 Spring Boot 3 为核心，辅以异步消息队列与 RESTful API，完成模型调度、推理计算及业务逻辑处理。检测算法方面，在 YOLOv8 基线模型上引入 SE (Squeeze-and-Excitation) 注意力机制，形成 YOLOv8-SE 轻量网络，有效适配边缘设备部署场景。

2.2 YOLOv8 模型基础

YOLOv8 是一种基于深度学习的单阶段目标检测模型，其核心架构由 Backbone-Neck-Head 三段式全卷积网络组成，通过优化特征提取、多尺度融合与任务解耦设计，实现了高精度与实时性的平衡。在特征提取阶段，YOLOv8 的 Backbone 基于改进的 CSPDarknet 架构，引入 C2f 模块作为核心组件，将输入特征沿通道维度拆分为两条路径：一条通过轻量化 Bottleneck 残差组提取细粒度纹理特征，另一条保留高维语义信息，最终通过 1×1 卷积重聚后融合，实现梯度复用与参数压缩。Backbone 集成空间金字塔池化模块，通过并行最大池化扩大感受野，同时将计算量降低 30%，显著提升小目标检测能力。在此基础上，YOLOv8 的 Neck 采用双向金字塔结构，通过自上而下的语义增强与自下而上的细节回传机制，实现多尺度特征融合。两路特征在同一尺度进行逐像素融合，并通过 C2f 模块替代传统卷积，增强跨尺度特征交互能力，使模型能够高效捕捉矿石目标在不同尺度下的特征。此外，YOLOv8

的 Head 模块摒弃传统锚框设计，采用 Anchor-Free 与解耦检测头：分类分支通过独立 1×1 卷积输出类别概率，利用 Varifocal Loss 优化分类置信度；回归分支通过独立 1×1 卷积直接预测边界框偏移量，结合 Distribution Focal Loss + CIoU Loss 提升定位精度。通过任务对齐匹配策略，模型在端到端训练中动态调整分类与回归损失权重，减少任务干扰，实现一次前向传播同步完成目标分类与定位。

2.3 SE 注意力机制

SE 注意力机制的核心模块主要包括以下几个方面：

2.3.1 压缩

对输入的特征图 $X \in \mathbb{R}^{C \times H \times W}$ 进行全局平均池化，将空间维度压缩为单个数值，得到通道描述向量 $z \in \mathbb{R}^C$ ：

$$z_j = \frac{1}{H \cdot W} \sum_{i=1}^{H \cdot W} x_j(i), \quad j = 1, \dots, C$$

该操作保留了每个通道的全局语义信息，同时显著降低了计算复杂度。

2.3.2 激励

通过两层全连接网络 (Fully Connected Network, FCN) 学习通道依赖关系，生成通道权重 $s \in [0, 1]^C$ ：

$$S = \sigma(W_2 \cdot \text{ReLU}(W_1 \cdot z))$$

其中 $W_1 \in \mathbb{R}^{C/r \times C}$ ， $W_2 \in \mathbb{R}^{C \times C/r}$ ， r 为压缩比。第一层全连接层将通道维度压缩至 C/r ，第二层恢复至原始维度，并通过 Sigmoid 函数生成归一化的通道权重。

2.3.3 重标定

将权重与原始特征图逐通道相乘，完成特征增强：

$$Y = s \otimes X$$

该操作动态放大重要特征 (如边缘、纹理) 的响应，抑制冗余通道，从而提升模型对关键信息的感知能力。

2.4 SE 模块在 YOLOv8 中的集成策略

YOLOv8 主干网络 (Backbone) 的核心构建单元是 C2f 模块。该模块通过引入多分支卷积结构和残差连接，高效地提取图像的基础特征，为后续的目标定位和分类任务提供丰富的语义信息。然而，尽管 C2f 模块在设计上兼顾了特征提取的广度和深度，但其内在机制存在一个显著的局限性：它缺乏对特征通道间依赖关系的显式建模。

为了克服 C2f 模块在特征提取方面的局限性，提升 YOLOv8 在复杂场景下的检测精度，引入了 SE 注意力机制。SE 核心思想在于显式地建模特征通道之间的相互依赖关系，从而让网络能够“有选择地”关注重要的特征，并抑制不重要的特征。通过将 SE 模块集成到 YOLOv8 的网

络结构中，可以显著增强模型的特征表达能力，使其能够更有效地处理复杂场景中的各种挑战，从而在不显著增加计算负担的前提下，实现检测精度的提升。

经过分析和实验验证，将 SE 模块放置在 C2f 模块之后被证明是一种高效且有效的策略。C2f 模块作为主干网络中的基础特征提取单元，其输出是一系列包含了丰富但未经筛选的特征图，在 C2f 模块之后立即接入 SE 模块，可以对这些“原始”的特征图进行及时的“清理”和“强化”。SE 模块能够评估 C2f 模块输出的每个通道的重要性，并对其进行加权，从而为后续的网络层提供更高质量、更具判别力的特征输入。此外，在特征提取或融合的关键节点上，利用 SE 模块对 C2f 模块产生的特征进行优化，从而提升整个网络的性能。

3 实验与结果分析

3.1 数据集

从 THE MINERAL & GEMSTONE KINGDOM 官网采集矿物图像，该平台收录超 1000 种矿物与宝石的专业摄影图像，图像分辨率普遍 ≥ 3000 × 2000 像素，具备专业级色彩还原。使用图像处理技术对采集到的矿物图像数据进行预处理，分析数据质量与标注准确性，应用数据增强技术（如随机裁剪、旋转），提升训练数据的多样性，使用学习曲线监控训练过程中的模型收敛情况。

3.2 评价指标

在矿物质目标检测任务中，mAP（mean Average Precision）是评估模型性能的核心指标，能够综合衡量模型对多类别、多尺度目标的检测能力。矿物质检测常涉及多种矿物成分（如金矿、铁矿等），不同类别样本数量可能差异显著，而 mAP 通过计算每个类别的平均精度（AP），再对所有类别取均值，能够公平反映模型在各类别中的表现，避免单一指标（如准确率）因类别不平衡导致的偏差。其公式计算如下：

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i$$

其中，N 为检测目标的类别总数，AP 为每个类别的平均精度，计算公式为：

$$AP_i = \int_0^1 \text{Precision}(r) dr$$

其中，Precision(r) 为在召回率（Recall）为 r 时的精度（正确检测的样本数 / 总检测数）。Recall 为正确检测的样本数 / 真实样本总数。

3.3 实验环境

表1 硬件与软件环境配置

环境类型	具体配置
训练平台	NVIDIA RTX 3090（24GB VRAM），Intel i9-11900K
推理平台	NVIDIA Jetson AGX Xavier（32GB RAM）
深度学习框架	PyTorch 1.12，CUDA 11.4

3.4 对比实验分析

为验证所提 YOLOv8+SE 注意力机制模型在矿物识别任务中的有效性与先进性，实验选取了目标检测领域主流的 Faster R-CNN、EfficientDet-D4、YOLOv7 模型与所提模型进行对比，在相同矿物数据集和实验环境下评估准确率、mAP@0.5:0.95 及推理速度指标。结果显示，所提 YOLOv8+SE 模型准确率达 95.2%，显著高于 Faster R-CNN 的 82.5%、EfficientDet-D4 的 88.0% 和 YOLOv7 的 89.0%；mAP@0.5:0.95 为 66.8%，优于其他模型的 58.3%、65.0% 和 65.5%；推理速度以 60ms / 样本远超对比模型，定量结果如表 2 所示。

表2 不同目标检测模型在矿物识别任务中的性能对比

模型类型	准确率	mAP@0.5:0.95	推理速度（ms / 样本）
FasterR-CNN	82.5%	58.3%	500
EfficientDet-D4	88.0%	65.0%	100
YOLOv7	89.0%	65.5%	150
YOLOv8+SE	95.2%	66.8%	60

为验证 SE 注意力机制对 YOLOv8 模型在矿物识别任务中性能的具体提升作用，设计了消融实验，对比分析 YOLOv8 基线模型、基线 + 通道注意力模块、基线 + 空间注意力模块以及基线 + SE 注意力机制四种方案在 30 类矿物识别任务中的表现，重点评估准确率、mAP@0.5 及 mAP@0.5:0.95 三个核心指标。实验结果显示，单独引入通道注意力模块后，准确率提升至 92.3%，mAP@0.5:0.95 提升至 65.2%；单独引入空间注意力模块后，准确率为 91.8%，mAP@0.5:0.95 为 64.8%；而融合 SE 注意力机制后，模型性能得到显著优化，准确率达到 95.2%，mAP@0.5 提升至 87.4%，mAP@0.5:0.95 提升至 66.8%，均优于单独引入通道或空间注意力模块的效果，定量结果如表 3 所示。

表3 SE注意力机制及组件消融实验结果

实验组	准确率	mAP@0.5	mAP@0.5:0.95
YOLOv8 基线模型	88.7%	85.3%	64.1%
基线 + 通道注意力模块	92.3%	86.5%	65.2%
基线 + 空间注意力模块	91.8%	86.0%	64.8%
基线 + SE 注意力机制	95.2%	87.4%	66.8%

在与 Faster R-CNN、EfficientDet-D4、YOLOv7 等主流目标检测模型的对比中，YOLOv8+SE 模型以

95.2% 的准确率显著领先于其他模型 (82.5%–89.0%)，mAP@0.5:0.95 达到 66.8%，高于 Faster R-CNN 的 58.3%、EfficientDet-D4 的 65.0% 和 YOLOv7 的 65.5%，同时推理速度以 60ms/ 样本远超对比模型，实现了“高精度 – 高速度”的双重突破。消融实验进一步揭示了 SE 注意力机制的增效机理：相较于 YOLOv8 基线模型（准确率 88.7%，mAP@0.5:0.95 64.1%），单独引入通道注意力或空间注意力模块可使准确率提升至 91.8%–92.3%，mAP@0.5:0.95 提升至 64.8%–65.2%；而融合 SE 注意力机制后，模型准确率提升至 95.2%，mAP@0.5 达 87.4%，mAP@0.5:0.95 达 66.8%，证明 SE 机制通过通道与空间特征的协同增强，有效提升了模型对矿物关键特征的捕捉能力，为矿物识别任务提供了更优的特征提取方案。

3.5 结果与讨论

3.5.1 所提模型在矿物识别准确率上显著优于主流目标检测模型

实验结果显示，所提 YOLOv8+SE 模型在 30 类矿物识别任务中准确率达到 95.2%，显著高于 Faster R-CNN 的 82.5%、EfficientDet-D4 的 88.0% 和 YOLOv7 的 89.0%，证明其在矿物细粒度识别任务中具备更强的特征判别能力。

3.5.2 引入 SE 注意力机制后模型推理速度较基线及其他模型大幅提升

引入 SE 注意力机制的 YOLOv8 模型边缘推理速度达 60ms/ 样本，较 YOLOv8 基线模型的 4000ms/ 样本提升 67 倍，且远超 Faster R-CNN 的 500ms/ 样本、EfficientDet-D4 的 100ms/ 样本和 YOLOv7 的 150ms/ 样本，满足实时检测需求。

3.5.3 消融实验验证通道与空间注意力模块对性能提升的协同作用

消融实验结果清晰显示，单独引入通道注意力模块时模型准确率提升至 92.3%，单独引入空间注意力模块时准确率为 91.8%，两者单独作用的性能提升幅度均有限；而当融合 SE 注意力机制实现通道与空间注意力的协同工作后，模型准确率显著跃升至 95.2%，且 mAP@0.5:0.95 达到

66.8%，远高于单一模块作用效果，充分验证了通道注意力对关键特征通道的筛选能力与空间注意力对局部细节的聚焦能力形成的协同增益，是模型性能提升的核心原因。

4 结语

本文针对矿物质检测场景中目标密集、背景纹理复杂、低层特征易受冗余干扰等难题，提出在 YOLOv8 主干网络 C2f 模块后引入 SE 通道注意力机制。该模块以 0.1 % 的参数量增量与 0.1 % 的推理延迟代价，对低层特征进行全局一通道重标定，显著提升了边缘、纹理等关键特征的表达能力。在自建矿物质数据集上的实验结果表明，所提方法的 mAP@0.5 由 85.3 % 提高到 87.4 %，mAP@0.5:0.95 由 64.1 % 提高到 66.8 %，小目标 AP 提升 2.3 %，验证了 SE 模块在矿物检测任务中的有效性。未来，将在轻量化部署与多光谱数据融合方向进一步拓展，以期对矿产资源智能勘探与实时监测提供更高效、可靠的解决方案。

参考文献：

- [1] 郭娟, 崔荣国, 周起忠等. 2023 年中国矿产资源形势回顾与展望[J]. 中国矿业, 2024,33(1):12–19.
- [2] 乔伊·雷等. YOLOv5：一种更快、更强的目标检测系统[J]. 计算机工程与应用, 2022(12):3001–3010.
- [3] 肖鹏程, 徐文广, 张妍等. 基于 SE 注意力机制的废钢分类评级方法[J]. 工程科学学报, 2023,45(8):1342–1352.
- [4] 蔡嘉磊, 茅智慧, 李君等. 基于深度学习的目标检测算法与应用综述[J]. 网络安全技术与应用, 2023(11):41–45.
- [5] 祁宣豪, 智敏, 图像处理中注意力机制综述[J]. 计算机科学与探索, 2024,18(2):346–349.

基金项目：广东理工学院科技项目：人工智能（2024KDZK001）；基于卷积神经网络的钢铁缺陷智能检测研究（2025YBZK016）。

作者简介：时小芳（1991.01–），女，汉族，河南濮阳人，研究生硕士，讲师，广东理工学院专任教师、教研室主任，研究方向：计算机视觉、深度学习领域。