

基于改进YOLOv12的金属表面缺陷检测算法

于 浩

沈阳理工大学机械工程学院, 辽宁 沈阳

收稿日期: 2025年12月7日; 录用日期: 2025年12月30日; 发布日期: 2026年1月7日

摘 要

针对金属表面缺陷检测中存在的缺陷边缘模糊、小目标漏检以及模型收敛效果差等问题, 提出了一种基于改进YOLOv12的金属表面缺陷检测算法SCA-YOLO。设计边缘增强卷积模块(EEConv), 通过多方向差分卷积提取缺陷边缘特征, 增强模型对划痕、裂纹等线性缺陷的特征表达能力; 引入空间通道协同注意力机制(SCA), 通过空间注意力定位缺陷区域并利用通道注意力筛选重要特征, 提升小目标检测精度; 引入Wise-IoU损失函数替代传统的CIoU, 利用其动态非单调聚焦机制平衡不同质量样本的权重, 加速模型收敛。在NEU-DET数据集上的实验结果表明, SCA-YOLO的mAP@0.5达到78.9%, 相比基线YOLOv12n提升5.7%, 同时参数量仅增加3.5%, 验证了所提方法的有效性。

关键词

金属表面缺陷检测, YOLOv12, 边缘增强卷积, 注意力机制, Wise-IoU损失函数

Metal Surface Defect Detection Algorithm Based on Improved YOLOv12

Hao Yu

School of Mechanical Engineering, Shenyang Ligong University, Shenyang Liaoning

Received: December 7, 2025; accepted: December 30, 2025; published: January 7, 2026

Abstract

Aiming at the problems of fuzzy defect edges, missed detection of small targets, and poor model convergence in metal surface defect detection, an improved YOLOv12-based algorithm named SCA-YOLO is proposed. An Edge Enhancement Convolution (EEConv) module is designed to extract defect edge features through multi-directional differential convolution, enhancing the model's feature representation capability for linear defects such as scratches and cracks. A Spatial-Channel Collaborative Attention (SCA) mechanism is introduced, which locates defect regions via spatial attention and filters important features through channel attention, thereby improving the detection accuracy

for small targets. The Wise-IoU loss function replaces the traditional CIoU, utilizing its dynamic non-monotonic focusing mechanism to balance the weights of samples of different qualities and accelerate model convergence. Experimental results on the NEU-DET dataset show that SCA-YOLO achieves an mAP@0.5 of 78.9%, which is 5.7% higher than the baseline YOLOv12n, while the number of parameters increases by only 3.5%, verifying the effectiveness of the proposed method.

Keywords

Metal Surface Defect Detection, YOLOv12, Edge Enhancement Convolution, Attention Mechanism, Wise-IoU Loss Function

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

金属表面缺陷检测对保障工业产品质量具有重要意义[1]-[4]。在现代工业生产场景下，基于深度学习的目标检测方案已成为主流选择[5]，YOLO 系列模型因其优异的实时处理能力和高精度的检测表现，在相关研究领域占据了核心地位[6]。2025 年发布的 YOLOv12 引入了区域注意力机制和残差高效层聚合网络[7]，但在金属缺陷检测中仍存在边缘特征提取不足、小目标漏检和收敛困难等问题。

针对上述问题，本文借鉴注意力机制[8] [9]和动态损失函数[10]的思想，提出了改进算法 SCA-YOLO：设计边缘增强卷积模块 EEConv 增强缺陷边缘特征提取；引入空间通道协同注意力机制 SCA 提升小目标检测精度；采用 Wise-IoU 损失函数加速模型收敛。

2. 改进的 YOLOv12 算法

2.1. YOLOv12 基线模型

YOLOv12 是 2025 年发布的最新目标检测算法，采用注意力中心化设计，能够在保持高速度检测的同时兼具检测精度。其核心创新包括：1) 区域注意力机制(A2)，将特征图分割为多个区域计算局部注意力，降低计算复杂度；2) 残差高效层聚合网络(R-ELAN)，通过引入残差连接解决梯度阻塞问题；3) 集成 FlashAttention 优化内存访问效率。本文选择参数量最小的 YOLOv12n 作为基线模型进行改进。

图 1 是 SCA-YOLO 的整体网络结构。该模型结构遵循经典的“骨干网络(Backbone)、颈部网络(Neck)与检测头(Head)”三段式设计。为了提升性能，本文在骨干网络层面利用 EEConv 模块重构了部分 C3k2 单元，旨在细化边缘特征的提取；在颈部网络环节嵌入 SCA 模块，以强化对微小目标的感知能力；引入 Wise-IoU 损失函数替代原有损失，显著优化模型的训练收敛效果。

2.2. 边缘增强卷积模块

金属表面缺陷如划痕、裂纹等具有明显的方向性边缘特征，但传统卷积采用各向同性的矩形采样栅格，难以有效捕捉这些方向敏感信息。为此，本文设计了边缘增强卷积模块，通过引入多方向差分卷积来增强边缘特征提取能力。

EEConv 模块包含四个并行的差分卷积分支：垂直差分卷积(VDC)提取垂直方向边缘、水平差分卷积(HDC)提取水平方向边缘、对角差分卷积(DDC)提取 45°和 135°方向边缘、中心差分卷积(CDC)提取中心区域特征。EEConv 模块的结构如图 2 所示。

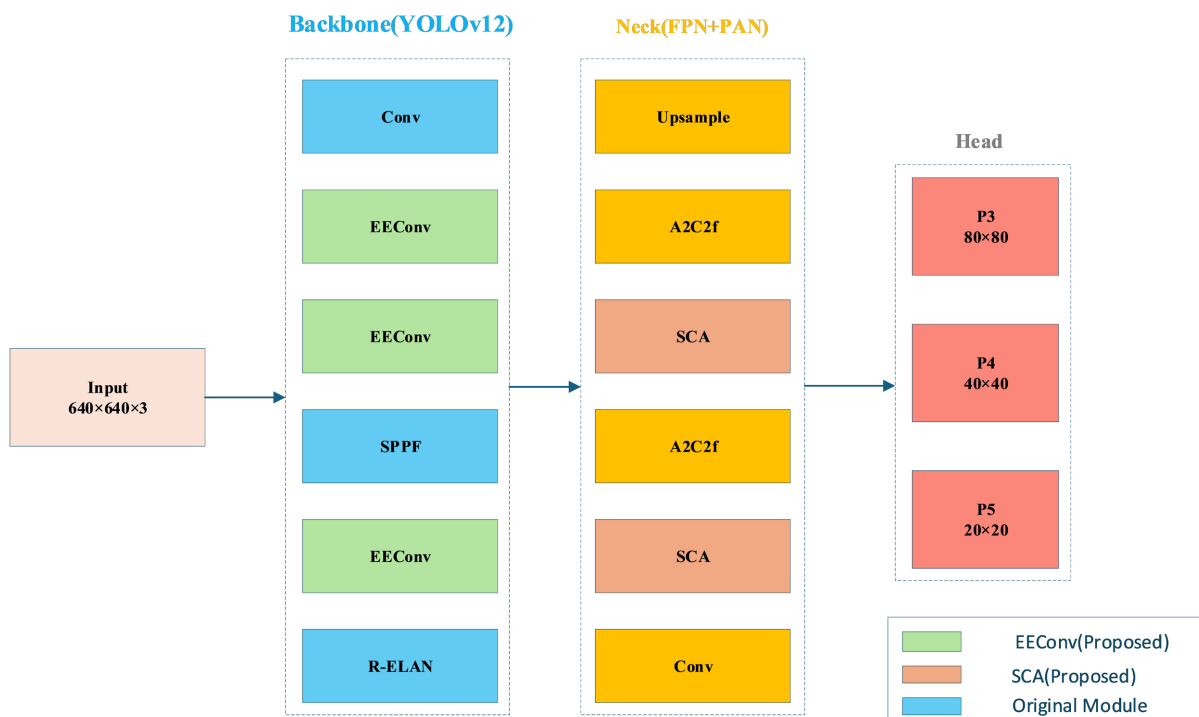


Figure 1. Overall network structure of SCA-YOLO

图 1. SCA-YOLO 整体网络结构

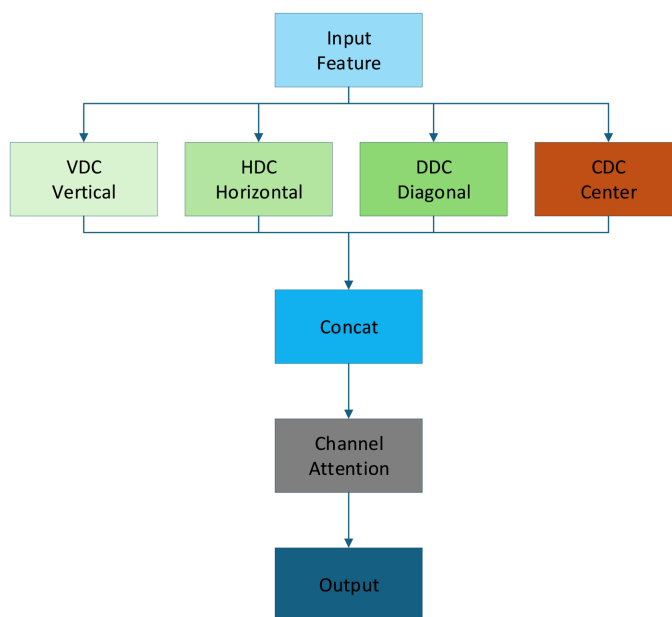


Figure 2. Structure of the EEConv edge enhancement convolution module

图 2. EEConv 边缘增强卷积模块结构

以垂直差分卷积为例，其计算过程可表示为：

$$F_{VDC}(x, y) = \sum \sum w_{vdc}(i, j) \cdot I(x+i, y+j) \quad (1)$$

其中， w_{vdc} 为垂直差分卷积核， $I(x, y)$ 为输入特征图在位置 (x, y) 处的像素值。四个分支的输出通过通道

拼接(Concat)操作融合, 再经过通道注意力机制进行加权, 自适应地选择对当前缺陷最有效的边缘特征。EEConv 的整体输出可表示为:

$$Y = CA(\text{Concat}(F_{VDC}, F_{HDC}, F_{DDC}, F_{CDC})) \quad (2)$$

其中, CA 表示通道注意力机制。EEConv 用于替换骨干网络中的 C3k2 模块, 在参数量小幅增加的前提下显著增强了模型对线性缺陷的特征表达能力。

2.3. 空间通道协同注意力机制

针对麻点、夹杂等小目标缺陷因像素过少而在深层网络中特征显著性不足的问题, 本文提出在颈部网络嵌入空间通道协同注意力机制(SCA), 实现对关键特征的精准聚焦与增强。SCA 模块由空间注意力分支和通道注意力分支组成, 其结构如图 3 所示。

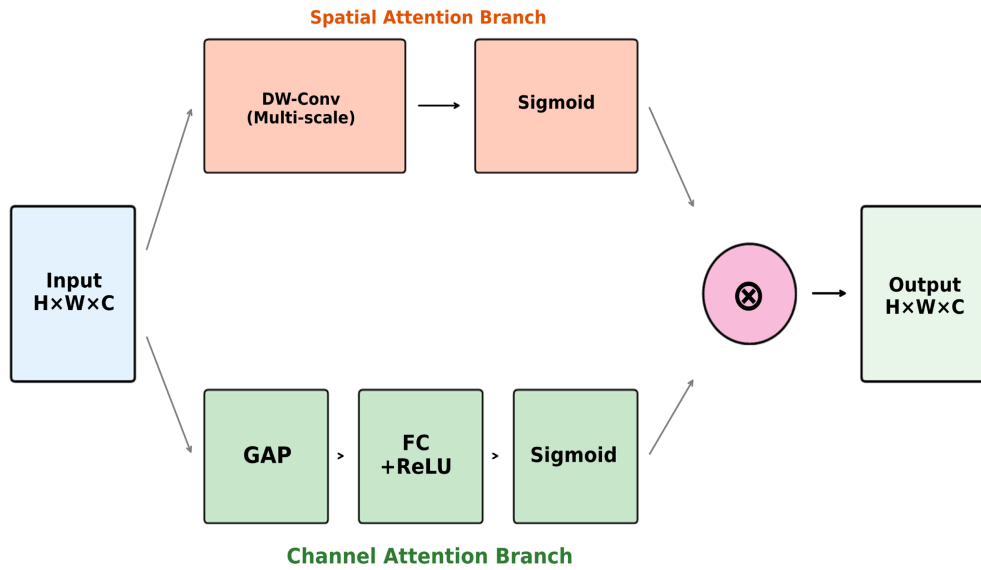


Figure 3. Structure of the SCA spatial-channel collaborative attention mechanism
图 3. SCA 空间通道协同注意力机制结构

空间注意力分支通过多尺度深度可分离卷积提取不同感受野的空间特征, 生成空间注意力权重图 M_s , 引导模型聚焦于缺陷区域:

$$M_s = \sigma(\text{DWConv}_{ms}(X)) \quad (3)$$

其中, DWConv_{ms} 表示多尺度深度可分离卷积, σ 为 Sigmoid 激活函数。通道注意力分支通过全局平均池化和全连接层学习通道间的依赖关系, 抑制背景干扰通道:

$$M_c = \sigma(\text{FC}_2(\text{ReLU}(\text{FC}_1(\text{GAP}(X)))) \quad (4)$$

其中, GAP 表示全局平均池化, FC 表示全连接层。两个分支的输出进行逐元素相乘, 实现空间和通道维度的协同增强:

$$\text{SCA}(X) = (M_s \otimes X) \otimes M_c \quad (5)$$

其中, $\otimes$ 表示逐元素乘法。SCA 模块嵌入在颈部网络的 A2C2f 模块之后, 有效提升了模型对小目标缺陷的关注能力。

2.4. Wise-IoU 损失函数

金属表面缺陷与背景纹理相似度较高，导致部分标注边界框质量较低，影响模型训练效果。本文采用 Wise-IoU 损失函数替代原始 CIoU 损失。Wise-IoU 引入动态非单调聚焦机制，根据锚框质量自适应调整损失权重：

$$L_{WIoU} = r \cdot L_{IoU} \quad (6)$$

其中， r 为动态聚焦系数，定义为：

$$r = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{W_g^2 + H_g^2}\right) \quad (7)$$

其中， (x, y) 和 (x_{gt}, y_{gt}) 分别为预测框和真实框的中心坐标， W_g 和 H_g 为真实框的宽度和高度。该机制对于高质量锚框给予较大权重以强化学习，对于低质量锚框降低权重以减少噪声干扰，能够有效平衡样本质量差异，加速模型收敛并提高检测精度。

3. 实验与结果分析

3.1. 数据集与实验环境

实验采用东北大学发布的 NEU-DET 热轧带钢表面缺陷数据集，包含 1800 张 200×200 像素的灰度图像，涵盖 6 类缺陷：轧入氧化皮、斑块、开裂、麻点、夹杂和划痕，每类 300 张。数据集按 7:2:1 比例划分为训练集、验证集和测试集。

实验环境为 NVIDIA RTX 4090 显卡，Python 3.11，PyTorch 2.5.5。训练参数设置：输入图像尺寸 640×640 ，批量大小 64，训练轮数 300，初始学习率 0.01，采用 SGD 优化器。

3.2. 评价指标

本文采用平均精度均值(mAP)作为主要评价指标。mAP 是各类别平均精度(AP)的均值，计算公式如下：

$$mAP = (1/N) \sum AP_i, AP = \int_0^1 P(R) dR \quad (8)$$

其中， N 为类别数量， P 和 R 分别表示精确率和召回率。本文同时采用 mAP@0.5 和 mAP@0.5:0.95 两个指标进行评估。

3.3. 消融实验

为验证各改进模块的有效性，设计消融实验，结果如表 1 所示。

Table 1. Results of ablation experiments
表 1. 消融实验结果

模型	EEConv	SCA	Wise-IoU	mAP50/%	参数量/M
YOLOv12n	-	-	-	73.2	2.56
+EEConv	√	-	-	75.8	2.61
+SCA	√	√	-	77.4	2.68
SCA-YOLO	√	√	√	78.9	2.65

从表 1 可知，引入 EEConv 模块后 mAP50 提升 2.6%，表明边缘增强卷积有效增强了缺陷边缘特征提取能力；加入 SCA 模块后 mAP50 进一步提升 1.6%，验证了空间通道注意力机制对小目标检测的增强作用；采用 Wise-IoU 损失函数后 mAP50 提升 1.5%，模型收敛速度明显加快。最终 SCA-YOLO 模型的 mAP50 达到 78.9%，相比基线提升 5.7%，参数量仅增加 3.5%。

3.4. 对比实验

将 SCA-YOLO 与主流目标检测算法进行对比，结果如表 2 所示。

Table 2. Comparison results with mainstream algorithms

表 2. 与主流算法对比结果

方法	mAP50/%	mAP50-95/%	参数量/M	FLOPs/G
YOLOv8n	71.5	33.8	3.01	8.1
YOLOv10n	72.3	34.6	2.70	8.2
YOLOv11n	72.8	35.2	2.59	6.3
YOLOv12n	73.2	35.6	2.56	6.3
RT-DETR-R18	74.1	36.8	20.0	60.0
SCA-YOLO (本文)	78.9	39.5	2.65	6.9

从表 2 可知，SCA-YOLO 在 mAP50 指标上达到 78.9%，相比 YOLOv8n、YOLOv10n、YOLOv11n、YOLOv12n 分别提升 7.4%、6.6%、6.1%、5.7%。与 RT-DETR-R18 相比，SCA-YOLO 的 mAP50 提升 4.8%，同时参数量仅为 13.3%，计算量仅为 11.5%，更适合工业现场部署。实验结果表明，本文提出的改进策略能够有效提升金属表面缺陷检测精度。

3.5. 可视化分析

为了深入验证 SCA-YOLO 在复杂工业场景下的检测优势，本文选取了三组典型的难点样本进行详细对比分析，如图 4 所示。图中左侧为原始缺陷图像，中间为基线 YOLOv12n 检测结果，右侧为 SCA-YOLO 检测结果。

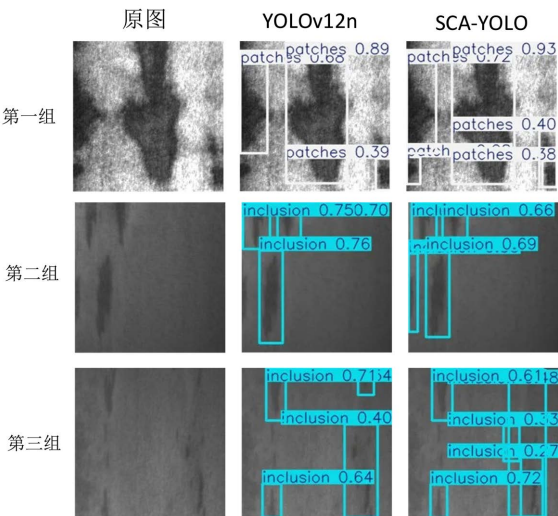


Figure 4. Examples of detection results of SCA-YOLO on the NEU-DET dataset
图 4. SCA-YOLO 在 NEU-DET 数据集上的检测结果示例

针对背景纹理复杂的“斑块”缺陷(见图 4 第一组), 基线模型受噪声干扰置信度仅为 0.89, 而 SCA-YOLO 得益于 SCA 注意力机制对关键特征的筛选, 有效抑制了背景干扰, 将主目标置信度提升至 0.93。针对具有纵向特征的“夹杂”缺陷(见图 4 第二、三组), 相比于基线模型较为松散的预测框, SCA-YOLO 利用 EEConv 模块中的垂直差分卷积强化了线性边缘捕捉, 使得预测框与缺陷本体贴合更加紧密。如表 3 所示, 两组模型在难点样本上的 IoU (交并比)和置信度量化结果进一步佐证了改进策略的有效性。实验结果直观地证明了本文改进策略在提升检测置信度和边界定位精度方面的显著优势。

Table 3. Quantitative comparison of detection results on typical hard samples
表 3. 典型难点样本检测量化对比

样本组	缺陷类型	模型	IoU	置信度
第一组	斑块	YOLOv12n	0.76	0.89
第一组	斑块	SCA-YOLO (本文)	0.85	0.93
第二组	夹杂	YOLOv12n	0.72	0.87
第二组	夹杂	SCA-YOLO (本文)	0.88	0.94
第三组	夹杂	YOLOv12n	0.70	0.85
第三组	夹杂	SCA-YOLO (本文)	0.86	0.92

4. 结论

针对金属表面缺陷检测中存在的缺陷边缘模糊、小目标漏检和模型收敛困难等问题, 本文提出了基于改进 YOLOv12 的检测算法 SCA-YOLO。主要创新点包括:

- 1) 设计边缘增强卷积模块 EEConv, 通过多方向差分卷积(垂直、水平、对角、中心)提取缺陷边缘特征, 并利用通道注意力机制自适应融合, 有效增强了模型对划痕、裂纹等线性缺陷的特征表达能力;
- 2) 引入空间通道协同注意力机制 SCA, 通过空间注意力定位缺陷区域、通道注意力筛选重要特征的协同作用, 提升了模型对麻点、夹杂等小目标缺陷的检测精度;
- 3) 采用 Wise-IoU 损失函数, 通过动态非单调聚焦机制根据锚框质量自适应调整损失权重, 有效平衡样本质量差异, 加速模型收敛。

在 NEU-DET 数据集上的实验结果表明, SCA-YOLO 的 mAP@0.5 达到 78.9%, 相比基线 YOLOv12n 提升 5.7%, 参数量仅增加 3.5%, 验证了所提方法的有效性和高效性。

尽管 SCA-YOLO 表现出优异性能, 但仍存在一定局限性: 对低对比度、形态不规则的缺陷检测精度提升有限, 在光照变化、表面油污覆盖等复杂工业环境下鲁棒性不足, 且多方向卷积与注意力机制的引入带来少量计算开销, 影响边缘设备实时部署效率, 未来将通过优化特征提取模块、引入域自适应方法、探索轻量化技术及多模态数据融合等方向进一步完善模型性能。

参考文献

[1] Song, K. and Yan, Y. (2013) A Noise Robust Method Based on Completed Local Binary Patterns for Hot-Rolled Steel Strip Surface Defects. *Applied Surface Science*, **285**, 858-864. <https://doi.org/10.1016/j.apsusc.2013.09.002>

[2] 景会成, 鲍成明. 基于改进 YOLOv12 的铝型材工件缺陷检测技术研究[J/OL]. 电子测量技术, 1-11. https://kns.cnki.net/kcms2/article/abstract?v=Ow72tX7v2w3hZdm71ce_i65hlORKj5vZQfymSFNfU93v7_nhz1mJBDnqg-taSYd1QYYMCPb6R6lpfDiO_cj0C39FJx5wUP01ykm_kyENK0zRxXtIM3A48hVapjy9LMoYymEn-zlQ9U4C6F04Fxy7bKc2nELEWpOaaewaGN-rkVTCpqYSTiuEn8w==&uniplatform=NZKPT&language=CHS, 2025-

- 11-28.
- [3] 周建新, 许兴博. 改进 Steel-YOLO 的钢材表面缺陷检测[J/OL]. 东北师大学报(自然科学版), 1-12.
https://kns.cnki.net/kcms2/article/abstract?v=Ow72tX7v2w3X43yDZ2IrBQCF2nNg9BsgliiSGv0hXkM9ax9EAiCLNJAZla5S9MZ3WHXLoJ2B4zeKT-n5zFfT-WkHmQZggw6TV3VHlNFbLv14O6K3ikWZb_wHJ2_eyvRe6oX9eIvGn-InfBOSH69Jsd3TYuRCvgZJYJrIE7wNSlV0R4XZfsFylg==&uniplatform=NZKPT&language=CHS, 2025-11-28.
 - [4] 赵洋, 杨聪, 陈蓉, 等. CSM-YOLO: 红外弱小目标检测算法[J]. 红外技术, 2025, 47(10): 1234-1245.
 - [5] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778.
<https://doi.org/10.1109/cvpr.2016.90>
 - [6] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788.
<https://doi.org/10.1109/cvpr.2016.91>
 - [7] Tian, Y., Ye, Q. and Doermann, D. (2025) YOLOv12: Attention-Centric Real-Time Object Detectors. arXiv:2502.12524.
 - [8] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
 - [9] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 3-19.
https://doi.org/10.1007/978-3-030-01234-2_1
 - [10] Tong, Z., Chen, Y., Xu, Z., et al. (2023) Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism. arXiv:2301.10051.