光电技术及应用

DOI: 10.16818/j.issn1001－5868.20251002001

基于改进YOLOv11n的轴承表面缺陷检测方法

周建新[1,2],焦江涛[1,2*],王 彬[1,2]

(1. 华北理工大学 电气工程学院,河北 唐山 063210;
2. 河北省风光氢储安全监测与智能运行技术创新中心,河北 唐山 063210)

摘 要: 针对轴承表面缺陷检测中存在的检测精度不足和模型效率不高等问题,文章提出一种改进YOLOv11n的缺陷检测模型GSP-YOLOv11n。首先,引入级联组注意力机制CGA,增强多头自注意力的上下文建模能力,提高对复杂纹理缺陷的感知精度。其次,引入Slim-Neck轻量化特征融合结构,在保证多尺度特征有效传递的同时,显著降低了计算复杂度。最后,引入一种分裂式空间压缩卷积结构SPConv,通过压缩冗余特征图,不仅有效减少了模型的计算量,还减少了参数规模,从而整体提升了模型的计算效率与轻量化程度。在自建轴承缺陷数据集上的实验结果表明,与原始YOLOv11n相比,GSP-YOLOv11n的*mAP*@0.5和*mAP*@0.5:0.95分别提升了2.9%和2.0%,参数量减少约9.7%,计算量降低1.2 G。该模型在轴承缺陷检测任务中表现优异,尤其对小目标和复杂形状缺陷的检测具有显著优势。

关键词: YOLOv11n;注意力机制;特征融合;空间压缩卷积;缺陷检测

中图分类号: TP391.41;TH133.3 **文章编号:** 1001－5868(2026)02－0380－08

Bearing Surface Defect Detection Method Based on Improved YOLOv11n

ZHOU Jianxin[1,2], JIAO Jiangtao[1,2], WANG Bin[1,2]

(1. College of Electrical Engineering, North China University of Science and Technology, Tangshan 063210, CHN;
2. Hebei Wind-Solar Hydrogen-Storage System Security Monitoring and Intelligent Operation Center of Technology Innovation, Tangshan 063210, CHN)

Abstract: To mitigate drawbacks such as insufficient detection accuracy and low model efficiency in bearing surface defect detection, a defect detection model, GSP-YOLOv11n, based on improved YOLOv11n is proposed. First, the cascade group attention mechanism is introduced to enhance the context modeling ability of multihead self-attention and improve the perception accuracy of complex texture defects. Second, the Slim-Neck lightweight feature fusion structure is incorporated, which ensures the effective transfer of multiscale features while significantly reducing the computational complexity. Finally, a split spatial compression convolution structure SPConv is introduced. Compressing redundant feature maps effectively reduces the computational cost of the model and the parameter scale, thus improving the overall computational efficiency and lightweight degree of the model. Experiments on a self-built bearing defect dataset demonstrate that GSP-YOLOv11n outperforms the original YOLOv11n model in terms of *mAP*@0.5 (2.9% enhancement), *mAP*@0.5:0.95 (2.0% enhancement), the number of parameters (9.7% reduction), and the computational cost (1.2 G reduction). Thus, the proposed model shows high bearing-defect-detection performance,

收稿日期: 2025－10－02.
基金项目: 河北省自然科学基金项目(F2018209201).
* **通信作者:** 焦江涛

offering significant advantages for the detection of small targets and defects with complex shapes.

Key words: YOLOv11n; attention mechanism; feature fusion; spatial compression convolution; defect detection

0 引言

轴承作为机械设备的核心部件，其表面缺陷（如凹槽、裂纹、划痕等）通常具有尺度微小、形态不规则及分布随机等特点。传统检测方法难以应对这些复杂的缺陷，因此，利用深度学习技术进行高效、准确的缺陷检测，成为提升轴承质量控制和保障设备安全运行的关键。

当前目标检测中的深度学习方法主要分为两类：单阶段目标检测算法和双阶段目标检测算法。其中，单阶段的目标检测算法主要有YOLO[1]，SSD[2]及Retina-Net[3]等；双阶段的目标检测算法主要有R-CNN[4]，Fast R-CNN[5]，Faster R-CNN[6]，SPP-Net[7]及Mask R-CNN[8]等。Wang等[9]利用灰度直方图成功实现了轴承缺陷的识别，然而，该方法中设计简洁的图像特征向量通常较为耗时。Liu等[10]开发了一套轴承内圈锥度和表面缺陷检测系统，使用工业相机采集图像并通过图像处理方法检测缺陷，尽管能有效检测锈蚀、划痕和污渍，但在小缺陷的检测上表现较差。Zhou等[11]通过多工位高效采集轴承表面图像，并设计图像预处理、区域提取和缺陷识别算法，准确识别了多种缺陷，但需要为每个表面设计不同的检测算法，流程较为复杂。高立明等[12]提出一种基于改进Faster R-CNN的绝缘轴承表面缺陷检测方法，该方法采用K-means++进行Anchor聚类，使用ROI Align替代ROI Pooling以避免量化误差，并通过在线难例挖掘策略提升困难样本的检测精度，但与单阶段目标检测算法相比，其检测速度仍有不足。李亚东等[13]在YOLOX骨干网络中引入Res2Net和上下文注意力机制，使用BiFPN代替PANet并加入坐标注意力，并通过Focal Loss增强模型对正样本的学习，尽管检测精度有提升，但导致了参数量增加，速度下降。Xu等[14]将YOLOv5骨干网中的C3模块替换为C2f模块，并在骨干和颈部网络中增加SPD模块，用CARAFE算子代替最近邻上采样，显著提升了凹痕和黑点等缺陷的检测精度，但在复杂场景下的检测仍需改进。姚景丽等[15]提出一种改进YOLOv8的轻量级轴承缺陷检测算法YOLO-SSW，该方法引入3-D注意力机制SimAM增强特征提取，在颈部嵌入C2f-SCConv模块以减少参数，同时加入小目标检测层提升小目标检测能力，并通过WIoU损失函数加速收敛，但计算量较大。王海群等[16]提出一种改进YOLOv10n的滚动轴承表面缺陷检测算法。该方法通过引入GhostConv，MSMHSA和CGLU模块重新设计C2f结构，构建CGMC2f模块以增强特征提取并减少参数量，在Neck部分，结合BIFPN，MAF-YOLO和EMCAD模块构建EMBS-FPN网络，提高检测精度并实现轻量化，但计算量仍需优化。魏梓书等[17]提出一种基于SBSI-YOLO11的轻量化轴承外观缺陷检测算法，通过引入SPD-Conv模块增强特征提取，采用BiFPN和SGE注意力机制提升小目标检测能力，并引入Inner-CIoU损失函数提高定位精度。该算法在精度和效率上有所提升，但计算量仍较大。

尽管现有方法在缺陷检测上有所提升，但仍难以满足实际需求。为此，本文提出一种基于改进YOLOv11n的轴承表面缺陷检测方法，具体包括：（1）引入级联组注意力机制CGA[18]，增强多尺度特征融合与小目标感知能力；（2）重新设计颈部网络，结合GSConv与VoV-GSCSP构建Slim-Neck[19]架构，在保证多尺度特征有效传递的同时显著降低计算复杂度；（3）引入SPConv[20]结构，通过将特征图划分为代表性和冗余部分，减少参数量和计算量，提升模型效率。

1 YOLOv11n的模型改进

1.1 YOLOv11n网络结构

YOLOv11n作为本研究的基础网络，在结构设计上进行了多项优化，整体由Backbone，Neck和Head三部分构成，结构如图1所示。Backbone部分引入了C3k2模块，该模块在保留C2f结构外观的同时，将内部瓶颈结构替换为C3K模块，增强了局部特征提取能力；在SPPF模块之后，新增C2PSA注意力模块，通过分层注意力机制提升对关键区域的感知能力，抑制无效特征干扰。Neck部分同样采用C3k2模块实现多尺度特征融合，有效提升对不同尺寸目标的检测性能。Head部分采用解耦式设计，分别对分类与回归任务进行建模，在降低计算开销的

同时保持准确性。YOLOv11n结构在检测效率和精度方面实现了良好的平衡，为本研究在后续轴承表面缺陷检测任务中进行结构改进奠定了基础。

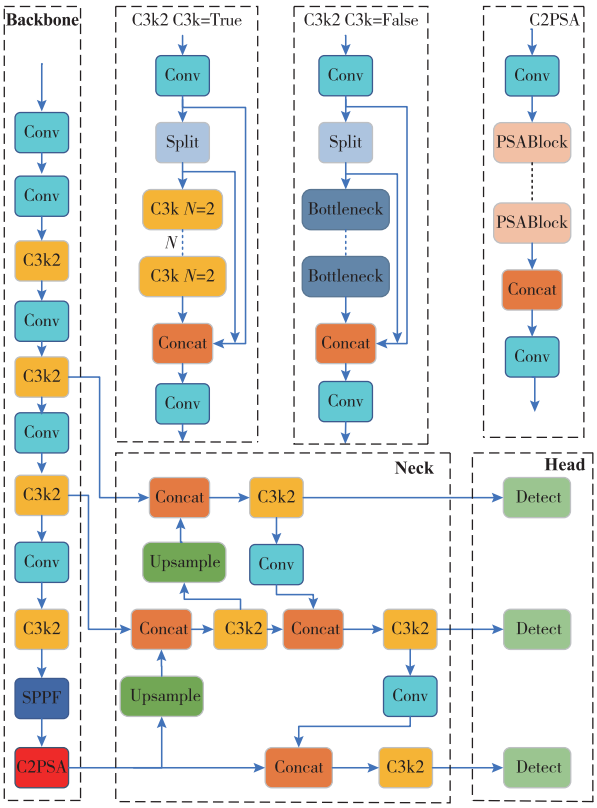


图1 YOLOv11n结构图

1.2 GSP-YOLOv11n网络结构

在轴承表面缺陷检测中，因缺陷形状复杂、纹理多变，现有检测器在检测精度与效率之间难以兼顾。为此，本文在YOLOv11n的基础上进行了改进，提出GSP-YOLOv11n模型，结构如图2所示。具体改进如下：在Neck中通过结合GSConv与VoV-GSCSP，构建Slim-Neck轻量化结构，对原YOLOv11n的Neck进行优化，在保证多尺度特征有效传递的同时显著降低计算复杂度；在每个VoV-GSCSP模块后引入了CGA模块，通过多尺度特征融合进一步增强上下文信息建模能力，从而提升模型对复杂缺陷纹理的感知精度，进而提高检测精度；在Backbone中使用分裂式空间压缩卷积（SPConv）替代原有部分卷积结构，有效压缩冗余特征图，提升小目标区域的表征能力，既减少了参数量，又降低了计算复杂度。GSP-YOLOv11n网络结构不仅显著提升了轴承缺陷检测精度，尤其在小目标缺陷较多的轴承数据集中表现优异，还有效减少了参数量和计算量，提高了计算效率，展现出更好的性能和效率平衡。

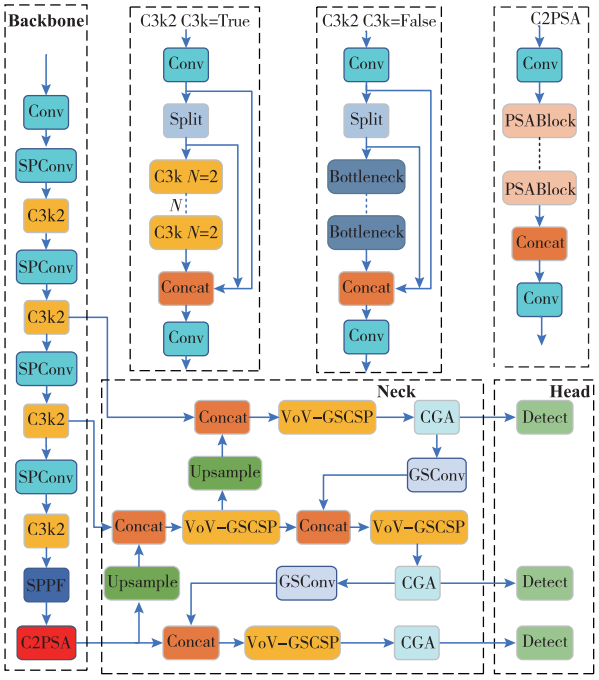


图2 GSP-YOLOv11n结构图

1.3 CGA结构模块

为增强模型在特征提取过程中的空间建模能力，在YOLOv11n结构中引入级联组注意力机制CGA，结构如图3所示。该模块源于EfficientViT的设计思想，旨在提升注意力机制的表达多样性。与传统多头自注意力机制（MHSA）不同，CGA通过将输入特征划分为多个子组，并为每个注意力头分配不同的特征子空间，实现了组内解耦。

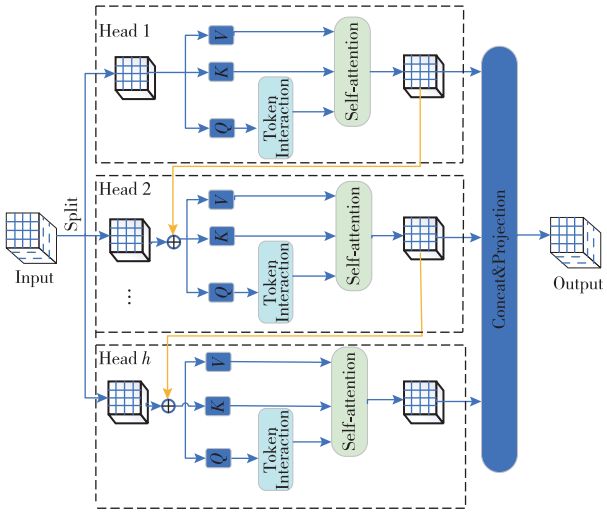


图3 CGA结构模块

对于每个注意力头，其输出由公式（1）计算：

$$\tilde{X}_{ij}=Attn\left(X_{ij}W_{ij}^{Q},X_{ij}W_{ij}^{K},X_{ij}W_{ij}^{V}\right) \quad (1)$$

式中，X_{ij}为第j个头的输入特征，$\boldsymbol{W}_{ij}^{Q}$，$\boldsymbol{W}_{ij}^{K}$，$\boldsymbol{W}_{ij}^{V}$分别为该头对应的查询、键、值映射矩阵。自注意力的计算采用缩放点积形式，即：

$$Attn(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=SoftMax\left(\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}/\sqrt{d}+\boldsymbol{B}\right)\boldsymbol{V} \quad (2)$$

式中，d为缩放系数，$\boldsymbol{B}$为可学习的位置偏置，用于引入相对位置信息。CGA 的核心在于注意力头之间的级联连接。除最后一头外，其余每一头的输入都会叠加前一头的输出，以增强信息流动性：

$$\boldsymbol{X}_{ij}'=\boldsymbol{X}_{ij}+\tilde{\boldsymbol{X}}_{i(j-1)},\ 2\leqslant j\leqslant h \quad (3)$$

最终，各头输出通过拼接并映射回原始维度，形成该层最终特征表示，即：

$$\tilde{\boldsymbol{X}}_{i+1}=Concat\left[\tilde{\boldsymbol{X}}_{ij}\right]_{j=1:h}\boldsymbol{W}_{i}^{p} \quad (4)$$

通过这种逐层增强的注意力结构，CGA 能更充分聚合轴承缺陷图像中的空间细节与上下文信息。

1.4　Slim-Neck 结构模块

为提升特征表达能力并降低模型计算开销，本研究引入了轻量化的 Slim-Neck 架构来改进 YOLOv11n 中的 Neck。Slim-Neck 通过用轻量级卷积 GSConv 替换传统卷积 Conv 实现模型轻量化，GSConv 如图 4 所示，结合了 SConv，DWConv，特征拼接（Concat）以及通道混洗（Shuffle）技术，能够高效提取特征，既降低了计算复杂度，又保留了模型的表现能力。在此基础上，原 C3k2 模块被 VoV-GSCSP 替换，采用单次跨阶段聚合策略，高效融合不同阶段的特征图，提升了特征利用效率并优化了计算成本，其结构如图 5 所示。实验结果表明，Slim-Neck 模块通过优化特征融合过程，减少了冗余计算，同时保持了高效的检测性能。

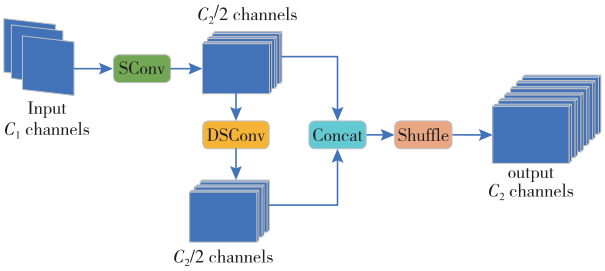


图 4　GSConv 模块

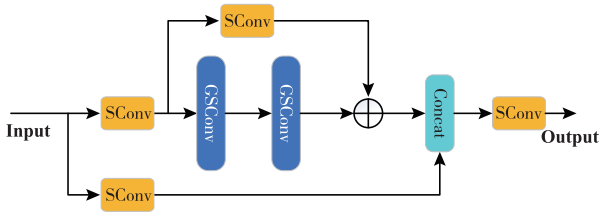


图 5　VoV-GSCSP 模块

1.5　SPConv 结构模块

上述改进创新虽然提高了检测精度并降低了计算量，但其代价是增加了模型的参数量。为了解决这一挑战，本研究引入了分裂式空间压缩卷积 SPConv 模块，替代主干网络中的 P2，P3，P4 和 P5 卷积层，有效缓解了常规卷积中的冗余特征问题。SPConv 结构通过分离冗余和代表性特征，提升了模型计算效率，具体结构如图 6 所示。

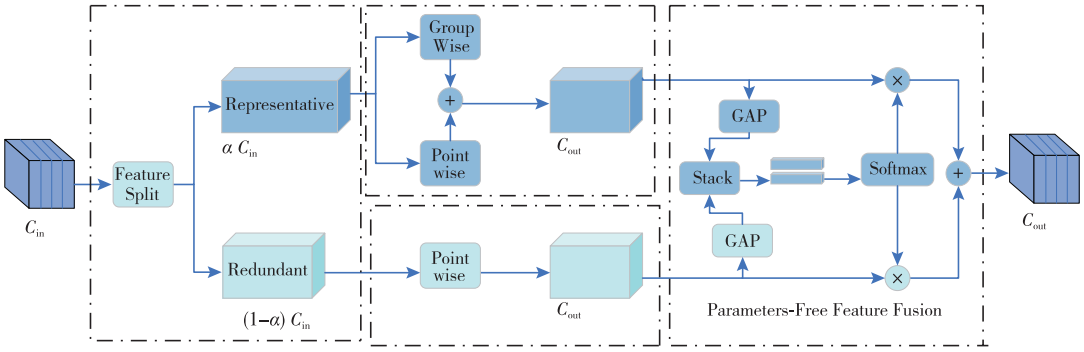


图 6　SPConv 结构模块

该结构的核心思想是将输入通道按照比例α划分为代表性特征与冗余特征两部分：代表性部分采用 3×3 卷积以提取关键信息，冗余部分则采用 1×1 卷积以补充细节，从而减少整体计算量。其拆分过程由公式（5）表示，W_{ij}表示 3×3 卷积核的参数，w_{ij}表示 1×1 卷积核的参数，x_i（$i=1,2,\cdots,L$）为输入特征矩阵，y_j（$j=1,2,\cdots,M$）为输出特征矩阵：

$$\begin{bmatrix}y_1\\y_2\\\vdots\\y_M\end{bmatrix}=\begin{bmatrix}W_{11}&\cdots&W_{1,aL}\\\vdots&\ddots&\vdots\\W_{M,1}&\cdots&W_{M,aL}\end{bmatrix}\begin{bmatrix}x_1\\\vdots\\x_{aL}\end{bmatrix}+\begin{bmatrix}w_{1,aL+1}&\cdots&w_{1,L}\\\vdots&\ddots&\vdots\\w_{M,aL+1}&\cdots&w_{M,L}\end{bmatrix}\begin{bmatrix}x_{aL+1}\\\vdots\\x_L\end{bmatrix} \quad (5)$$

将输入特征根据α划分为代表性和冗余部分后，针对代表性部分的冗余，通过3×3卷积对通道进行划分，减少冗余特征。为避免信息丢失，结合1×1点卷积恢复缺失信息，最终通过加法操作融合两者的卷积结果，如公式(6)所示：

$$\begin{bmatrix} W_{11}^{P} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & W_{GG}^{P} \end{bmatrix}\begin{bmatrix} z_1 \\ \vdots \\ z_G \end{bmatrix}+\begin{bmatrix} W_{11} & \cdots & W_{1,aL} \\ \vdots & \ddots & \vdots \\ W_{M,1} & \cdots & W_{M,aL} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_{aL} \end{bmatrix} \tag{6}$$

式中，W_{ll}^{P}表示第l组卷积核的参数，把αL个代表性通道分成G组，每组包含$\alpha L/G$个通道，$z_i(i=1,2,\cdots,G)$是第i组$\alpha L/G$通道经过组卷积操作后的输出结果。

通过上述操作得到代表性部分和冗余部分的特征输出。为了提取通道信息，首先对特征进行全局平均池化，生成统计量S_1和S_3，并将其相加后通过全连接层得到压缩量Z。接着，使用Softmax归一化处理，得到β和γ，并将其分别与代表性和冗余特征相乘。最终，通过加法操作融合两者，生成输出特征，整个过程如公式(7)所示：

$$\begin{aligned} Y=W'X\approx\beta&\left\{\begin{bmatrix} W_{11}^{P} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & W_{GG}^{P} \end{bmatrix}\begin{bmatrix} z_1 \\ \vdots \\ z_G \end{bmatrix}+\right. \\ &\left.\begin{bmatrix} W_{11} & \cdots & W_{1,aL} \\ \vdots & \ddots & \vdots \\ W_{M,1} & \cdots & W_{M,aL} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_{aL} \end{bmatrix}\right\}+ \\ &\gamma\begin{bmatrix} w_{1,aL+1} & \cdots & w_{1,L} \\ \vdots & \ddots & \vdots \\ w_{M,aL+1} & \cdots & w_{M,L} \end{bmatrix}\begin{bmatrix} x_{aL+1} \\ \vdots \\ x_L \end{bmatrix} \end{aligned} \tag{7}$$

常规卷积的参数量可以用公式(8)来表示：

$$P_{\text{Conv}}=k\times k\times L\times M \tag{8}$$

SPConv的参数量可以用公式(9)来表示：

$$\begin{aligned} P_{\text{SPConv}}=&k\times k\times\frac{1}{g}\alpha L\times\frac{1}{g}\times M\times g+1\times 1\times \\ &\alpha L\times M+1\times 1\times(1-\alpha)L\times M= \\ &\left(k\times k\times\frac{\alpha}{g}+1\right)\times L\times M \end{aligned} \tag{9}$$

式中，g为特征图像通道划分组数。

因此，SPConv通过替换主干网络中的部分常规卷积，采用组卷积提取形态和纹理特征，点卷积则整合通道信息并补充细节，从而有效消除特征层间的冗余并提高特征提取效率。SPConv不仅减少了参数量，还加快了网络推理速度。

2 实验数据集和实验环境

2.1 数据集的制作

通过爬虫技术从搜索引擎和公开资源抓取相关缺陷图像，并经过筛选和使用LabelImg工具对缺陷类别进行标注，最终形成自制轴承缺陷检测数据集。该数据集为单张图片同时存在多种缺陷的复合故障数据集，包含6类缺陷：铸造毛刺、抛光铸件、裂纹、凹坑、划痕和未抛光铸件。数据集共计3 659张图像，图像分辨率为640×640，并按7∶2∶1的比例划分为训练集(2 561张)、验证集和测试集。训练集中各类缺陷的样本数分别为：铸造毛刺536张，抛光铸件1 506张，裂纹1 154张，凹坑592张，划痕1 687张，未抛光铸件571张。部分数据集如图7所示。



图7 部分数据集

2.2 实验环境

系统选择Windows10，硬件环境包括NVIDIA GeForce RTX 4060(8 GB显存)显卡和第13代Intel Core i5-13400F处理器。采用PyTorch 2.0.0深度学习框架，Python版本为3.11.4，CUDA版本为11.8。训练轮数设为200，批量大小为16，使用8个线程加载数据。优化器选用SGD，初始学习率为0.01，动量设置为0.937。

3 实验结果与分析

3.1 评价指标

实验采用精确率(P)、召回率(R)、$mAP@0.5$和$mAP@0.5{:}0.95$作为主要评估指标。其中$mAP@0.5$表示，当IoU阈值为0.5时，各类别的平均精度；$mAP@0.5{:}0.95$则表示，当IoU阈值从0.5到0.95，步长为0.05时，计算的平均精度，用于更全面地评估模型检测效果。在模型轻量化方面，通过参数量(Parameters)和计算量(GFLOPs)来衡量模型的复杂度与计算效率，相关公式如下：

$$P=\frac{TP}{TP+FP} \tag{10}$$

$$R=\frac{TP}{TP+FN} \tag{11}$$

$$AP=\int_0^1 P(R)\mathrm{d}R \tag{12}$$

$$mAP=\frac{1}{n}\sum_{i=1}^{n}AP(i) \tag{13}$$

式中，TP为正确检测的正样本数，FP为误检的负样本数，FN为漏检的正样本数。mAP为所有类别AP的平均值，n为类别数，i为检测次数。

3.2　消融实验

为验证各改进模块的有效性，在YOLOv11n基线模型上逐步引入CGA，Slim-Neck和SPConv进行消融实验，结果如表1所示。

表1　消融实验结果

模型	*P*/%	*R*/%	*mAP*@0.5/%	*mAP*@0.5:0.95/%	*Parameters*/M	*GFLOPs*/G
YOLOv11n	79.5	84.5	85.0	65.7	2.58	6.3
YOLOv11n+CGA	79.6	85.2	85.9	66.1	2.71	6.6
YOLOv11n+Slim-Neck	80.1	84.6	85.7	65.7	2.60	5.8
YOLOv11n+SPConv	77.3	86.0	85.1	65.7	2.28	5.4
YOLOv11n+CGA+Slim-Neck	78.8	86.2	86.3	66.2	2.73	6.1
YOLOv11n+CGA+SPConv	79.2	86.3	86.8	67.0	2.40	5.7
YOLOv11n+Slim-Neck+SPConv	80.8	84.5	86.7	66.6	2.21	4.8
GSP-YOLOv11n	81.0	85.6	87.9	67.7	2.33	5.1

实验结果表明，单独引入CGA模块显著提升了召回率与mAP指标，有效增强了特征提取与目标感知能力。单独引入Slim-Neck模块，能够在降低计算复杂度并提升推理速度的同时，保持较高的检测精度。单独加入SPConv模块，通过压缩冗余特征，使参数量减少11.6%、计算量降低0.9 G，验证了其在轻量化与效率优化方面的有效性。

进一步的组合实验表明，各模块在性能提升上具有互补性：CGA与Slim-Neck结合后在mAP和R上均有所提高；CGA与SPConv结合在精度提升的同时显著降低了参数量与计算量；Slim-Neck与SPConv结合则在轻量化效果上更为突出，参数量和计算量均明显降低，同时保持了较高的精度表现；而GSP-YOLOv11n综合性能最佳，相较基线模型，P提升1.5%，R提升1.1%，mAP@0.5提升2.9%，mAP@0.5:0.95提升2.0%，同时参数量减少9.7%、计算量降低1.2 G，体现出精度与效率的协同优化优势。

3.3　不同模型对比实验

为验证所提方法的有效性，将GSP-YOLOv11n与SSD，Faster R-CNN及YOLO系列不同版本模型进行了对比，使用mAP@0.5，mAP@0.5:0.95，P，GFLOPs和Parameters作为性能指标对不同模型进行评估，实验结果如表2所示。

表2　不同模型对比实验结果

模型	*P*/%	*mAP*@0.5/%	*mAP*@0.5:0.95/%	*Parameters*/M	*GFLOPs*/G
SSD	77.3	82.4	61.8	24.41	138.19
Faster R-CNN	79.2	83.4	62.6	41.15	91.03
YOLOv3-tiny	79.4	83.2	62.7	12.13	18.9
YOLOv5n	77.1	84.8	65.0	2.50	7.1
YOLOv6n	77.5	84.1	64.6	4.23	11.8
YOLOv8n	79.4	84.9	65.2	3.00	8.1
YOLOv10n	76.7	82.9	64.0	2.26	6.5
YOLOv10s	79.3	83.8	65.0	7.22	21.4
YOLOv11n	79.5	85.0	65.7	2.58	6.3
GSP-YOLOv11n	81.0	87.9	67.7	2.33	5.1

结果表明，传统检测模型SSD与Faster R-CNN在检测精度和计算效率方面均表现不足，难以满足轻量与实时需求。而YOLO系列轻量模型在精度与效率间取得了较好平衡，其中YOLOv11n基线模型以85.0%的mAP@0.5，2.58 M参数量和6.3 G的计算量展现出优越的综合性能。相比之下，本文提出的GSP-YOLOv11n表现更加优异，在mAP@0.5提升至87.9%、mAP@0.95达67.7%的同时，参数量与计算量分别降至2.33 M与5.1 G，显著优于现有各类轻量模型，在保持高精度的同时实现了更优的推理效率。

3.4　不同卷积对比实验

为体现SPConv在YOLOv11n上的改进性能，在Backbone分别引入AKConv，DualConv，RFAConv，RefConv替代P2，P3，P4和P5卷积层进行对比试验，结果如表3所示。SPConv在保持较高检测精度的同时，实现最优轻量化效果，参数量与计算量分别仅为2.28 M和5.4 G，较基准Conv降低11.6%和14.3%。相比之下，AKConv(2.29 M，5.6 G)和DualConv(2.40 M，5.9 G)虽然也展现出一定的轻量化效果，但优化幅度均不及SPConv。值得注意的是，RFAConv(2.63 M，6.7 G)和RefConv(2.59 M，8.5 G)在引入后反而导致模型复杂度上升，特别是RefConv使计算量显著增加了34.9%。结果表明，SPConv在精度保持与模型轻量化方面具有显著优势。

表 3 不同卷积对比实验结果

模块	*R*/%	*mAP*@0.5/%	*Parameters*/M	*GFLOPs*/G
Conv	84.5	85.0	2.58	6.3
AKConv	85.8	84.2	2.29	5.6
DualConv	85.5	84.4	2.40	5.9
RFAConv	85.1	85.3	2.63	6.7
RefConv	84.3	85.2	2.59	8.5
SPConv	86.0	85.1	2.28	5.4

3.5 模型热力图分析

为评估改进后模型的特征提取效果，本研究对比了原始模型与改进模型的6种缺陷热力图，结果如图8所示，其中红色区域代表模型关注的重点特征区域。从图中可见，改进后的模型对缺陷区域关注更集中，热力区域面积更大，具备更强的特征表达和抗干扰能力。

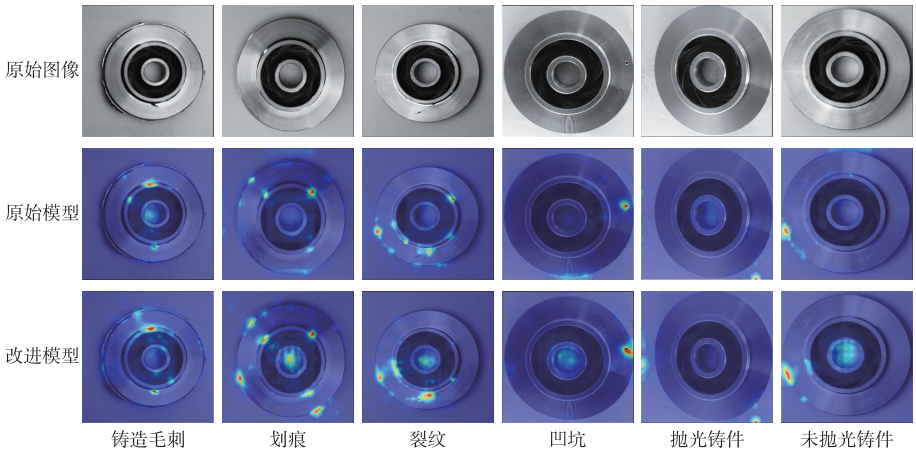
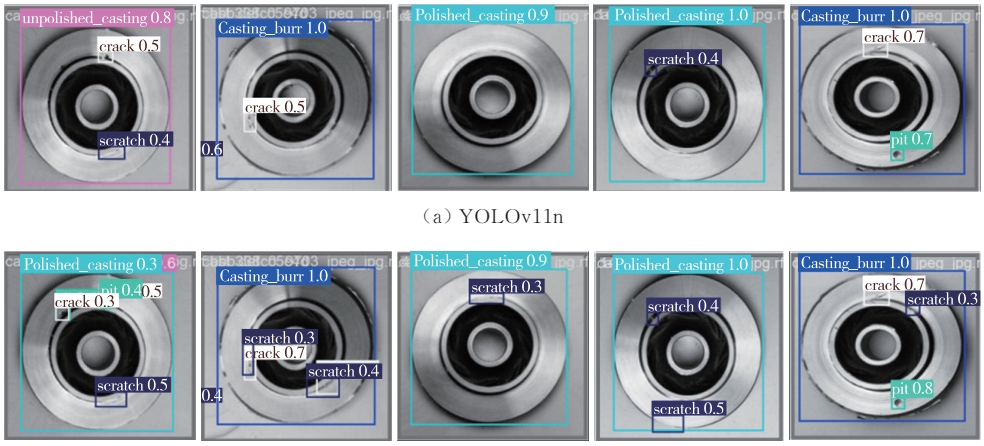


图 8 热力图可视化对比结果

3.6 可视化分析

为清晰展示改进后算法的检测效果，将GSP-YOLOv11n模型与YOLOv11n模型在自建轴承缺陷数据集上进行对比，使用不同颜色的锚框区分各缺陷类别。可视化对比结果如图9所示。通过对比改进前后的检测效果，YOLOv11n模型在检测裂纹和划痕等缺陷时存在误检和漏检。GSP-YOLOv11n模型显著提高了检测精度，减少了误检和漏检问题。结果表明，GSP-YOLOv11n在轴承表面缺陷检测中具有更好的性能。



（a）YOLOv11n

（b）GSP-YOLOv11n

图 9 可视化对比结果

4 结论

本研究针对轴承表面缺陷检测中存在的形状复杂、纹理多变及漏检率高等问题，提出一种GSP-YOLOv11n检测模型。该模型通过设计级联组注意力机制CGA，Slim-Neck轻量化结构以及SPConv模块，有效提升了其对复杂缺陷的感知能力，同时降低了计算复杂度。实验结果表

明，GSP-YOLOv11n在精度、召回率和检测性能方面均优于原YOLOv11n，且具备更高的计算效率。

参考文献：

[1] 徐彦威，李　军，董元方，等. YOLO系列目标检测算法综述[J]. 计算机科学与探索，2024，18(9)：2221-2238.
Xu Yanwei，Li Jun，Dong Yuanfang，et al. Survey of development of YOLO object detection algorithms[J]. Journal of Frontiers of Computer Science & Technology，2024，18(9)：2221-2238.

[2] Jiang T. Application of SSD network algorithm in panoramic video image vehicle detection system [J]. Open Computer Science，2023，13：20220270.

[3] Weikert T，Jaeger P F，Yang S，et al. Automated lung cancer assessment on 18F-PET/CT using Retina U-Net and anatomical region segmentation[J]. European Radiology，2023，33(6)：4270-4279.

[4] Girshick R，Donahue J，Darrell T，et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition，2014：580-587.

[5] Girshick R. Fast R-CNN [C]//2015 IEEE International Conference on Computer Vision，2016：1440-1448.

[6] Seo D M，Woo H J，Kim M S，et al. Identification of asbestos slates in buildings based on faster region-based convolutional neural network (faster R-CNN) and drone-based aerial imagery [J]. Drones，2022，6(8)：194.

[7] Wang X，Wang S，Cao J，et al. Data-driven based tiny-YOLOv3 method for front vehicle detection inducing SPP-net [J]. IEEE Access，2020，8：110227-110236.

[8] Vu T，Bao T，Hoang Q V，et al. Measuring blast fragmentation at Nui Phao open-pit mine，Vietnam using the Mask R-CNN deep learning model[J]. Mining Technology，2021，130(4)：232-243.

[9] Wang J，Qiao J，Guo M. Research on bearing surface defect detection system based on machine vision [J]. Journal of Physics：Conference Series，2022，2290(1)：012061.

[10] Liu C，Chen Z，Liu P，et al. Intelligent bearing quality checking system based on visual inspection[C]//2020 Chinese Automation Congress，2021：3819-3825.

[11] Zhou P，Zhang C，Zhou G，et al. Whole surface defect detection method for bearing rings based on machine vision [J]. Measurement Science and Technology，2023，34(1)：015017.

[12] 高立明，贾书海，张国龙，等. 基于改进Faster R-CNN的绝缘轴承表面缺陷检测方法[J]. 轴承，2023(4)：1-8.
Gao Liming，Jia Shuhai，Zhang Guolong，et al. Detection method for surface defects of insulated bearings based on improved faster R-CNN[J]. Bearing，2023(4)：1-8.

[13] 李亚东，马　行，穆春阳，等. 改进YOLOX网络的轴承缺陷小目标检测方法[J]. 计算机工程与应用，2023，59(1)：100-107.
Li Yadong，Ma Hang|Xing)，Mu Chunyang，et al. Improved small target detection method of bearing defects in YOLOX network[J]. Computer Engineering and Applications，2023，59(1)：100-107.

[14] Xu H，Pan H，Li J. Surface defect detection of bearing rings based on an improved YOLOv5 network[J]. Sensors，2023，23(17)：7443.

[15] 姚景丽，程　光，万　飞，等. 改进YOLOv8的轻量化轴承缺陷检测算法[J]. 计算机工程与应用，2024，60(21)：205-214.
Yao Jingli，Cheng Guang，Wan Fei，et al. Improved lightweight bearing defect detection algorithm of YOLOv8 [J]. Computer Engineering and Applications，2024，60(21)：205-214.

[16] 王海群，陈晓宇，于海峰. 基于改进YOLOv10n的滚动轴承表面缺陷检测 [J/OL]. 电子测量技术，1-12[2025-09-30].
Wang Haiqun，Chen Xiaoyu，Yu Haifeng. Detection of Surface Defects of Rolling Bearings Based on Improved YOLOv10n [J/OL]. Electronic measurement technology，1-12[2025-09-30].

[17] 魏梓书，陈志刚，王衍学，等. 基于SBSI-YOLO11的轻量化轴承外观缺陷检测算法[J/OL]. 广西师范大学学报(自然科学版)，1-12[2025-09-30].
Wei Zishu，Chen Zhigang，Wang Yanxue，et al. Lightweight Bearing Appearance Defect Detection Algorithm Based on SBSI-YOLO11[J/OL]. Journal of Guangxi Normal University (Natural Science Edition)，1-12 [2025-09-30].

[18] Liu X，Peng H，Zheng N，et al. EfficientViT：memory efficient vision transformer with cascaded group attention [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition，2023：14420-14430.

[19] Li H，Li J，Wei H，et al. Slim-neck by GSConv：a lightweight-design for real-time detector architectures [J]. Journal of Real-Time Image Processing，2024，21(3)：62.

[20] Zhang Q，Jiang Z，Lu Q，et al. Split to be slim：an overlooked redundancy in vanilla convolution [EB/OL]. 2020：arXiv：2006.12085.

作者简介：

周建新(1977—)，男，河北省唐山市人，副教授，硕士生导师，主要研究方向为智能控制理论及应用；

焦江涛(2002—)，男，河北省邯郸市人，硕士研究生，研究方向为智能控制与模式识别。