

# 主干特征提取与融合的连接铸坯表面缺陷检测轻量化模型研究

成彬, 井长鑫, 王井浩, 陆波

(西安建筑科技大学 机电工程学院, 陕西西安 710055)

**摘要:** 针对计算资源受限情况下连接铸坯表面缺陷检测算法在模型部署与实际应用中效率不高的问题, 本文提出了一种基于深度学习模型轻量化的连接铸坯表面缺陷检测方法, 旨在实现高效且快速的缺陷检测。首先, 设计了基于GhostNet的主干特征提取网络, 显著提高了特征提取的效率, 同时大幅降低了网络参数量; 其次, 构建了基于VoVGSCSP与GSConv模块的特征融合网络, 通过优化特征融合策略, 进一步减少了模型的内存占用与计算开销; 最后, 面向连接铸坯表面缺陷数据集和热轧带钢表面缺陷数据集进行对比试验。试验表明, 本文方法在多项性能指标上表现优异: 平均检测精确率达0.773, 检测速度提升至138.89帧/s; 模型的权重大小为8.3 MB, 参数量为 $3.91 \times 10^5$ , 浮点运算次数(FLOPs)降低至7.4 G。该方法相比目前主流模型在复杂度方面模型轻量化优势显著, 同时在检测速度和检测精度之间实现了良好平衡, 为资源受限环境下连接铸坯表面缺陷检测的深度学习算法研究提供了一种具有较高参考价值的新思路和技术实现方案。

**关键词:** 连接铸; 表面缺陷; 智能检测; 轻量化; 特征提取与融合

## 作者简介:

成彬(1972-), 男, 教授, 博士生导师, 研究方向为数字化智能化设计与制造、机器视觉与数字孪生。  
E-mail: cheengbin@163.com.

## 通信作者:

井长鑫, 男, 硕士生, E-mail: 2876344958@qq.com; 王井浩, 男, 硕士生, E-mail: 982424050@qq.com

中图分类号: TG249.7;  
TP391.4

文献标识码: A

文章编号: 1001-4977(2025)  
04-0500-08

## 收稿日期:

2024-12-18.

连接铸坯是连接铸工艺生产的重要中间产品, 此产品的表面质量对后续产品的整体表面属性及质量性能产生重要的影响。然而, 由于受物性参数、拉速及压下参数等多环境因素的干扰, 连接铸坯表面容易产生多种严重的表面缺陷, 而此类表面质量问题可能导致后续产品在轧制工艺或使用中会出现一系列严重的质量问题<sup>[1]</sup>。因而对连接铸坯表面缺陷进行实时检测和表面缺陷的总体统计及预测, 提升连接铸坯质量等具有重要的理论意义和工程应用价值。

目前, 针对连接铸坯表面产生的缺陷及其检测方法等问题, 国内外主要依赖人工经验检测, 通过抽样检查来控制其表面质量。然而, 人工经验检测存在较大的漏检与误检, 同时其检测的实时速度不能满足实际生产需要<sup>[2]</sup>。此外, 国内外学者应用基于人工智能的机器视觉、深度学习及其他检测技术实现表面缺陷检测。在国外, 阿美特克视觉检测公司(AMETEK Surface Vision)开发了连接铸坯机器视觉检测装置; 普锐特冶金技术公司(Primetals Technologies)应用X射线、超声波和磁粉检测等技术开发了面向多种冶金产品缺陷的检测系统; 欧洲实验室公司(TEC Eurolab)应用多领域超声波、磁粉检测和热成像等技术开发了无损缺陷装置。在国内, 欧阳奇运用脉冲电涡流检测技术对高温铸坯表面缺陷进行检测<sup>[3]</sup>。丁恒针对工业生产中铸坯内部质量控制的需求, 提出了基于分层聚焦扫描的超声显微无损检测方法<sup>[4]</sup>。张雪元采用面阵相机及激光扫描相结合, 对连接铸坯表面进行实时检测与分析研究<sup>[5]</sup>。徐科以线激光为光源利用线阵相机采集铸坯图像特征, 并以AdaBoosting算法对铸坯表面缺陷进行识别<sup>[6]</sup>。相比以上模型方法, 应用深度学习方法构建的检测模型对连接铸坯表面缺陷检测效果和精度提升等方面体现了更强的性能, 以及此模型具有更好的鲁棒性和普适性, 因此, 胡嘉成使用支持向量机和卷积神经网络对上海宝钢公司连接铸坯表面缺陷进行检测, 验证了对裂纹缺陷检测的有效性<sup>[7]</sup>。王睿泽在连接铸坯表面缺陷识别

任务中,基于YOLOv3算法进行轻量化改进,验证了对面部裂纹、角部划痕和面部划痕检测的有效性<sup>[8]</sup>。成彬等提出一种基于改进YOLOv5的连铸坯表面缺陷检测方法,有效解决了裂纹、气孔和夹杂三类缺陷的漏检误检及检测精度低等问题<sup>[9]</sup>。一些研究者通过引入注意力机制、多尺度融合<sup>[10-12]</sup>等模块改进结构来增强模型对缺陷的感知能力。

尽管国内外研究者通过引入机器视觉、无损检测技术和深度学习算法在连铸坯表面缺陷检测方面取得了一定进展,但这些方法仍存在检测速度慢、精度不足以及实时性差等问题,无法同时满足工业生产中对高效和高精度检测的需求。

因此,本文基于深度学习技术,以平衡检测精度与检测速度为研究目的,基于GhostNet轻量化思想实现主干网络构建以提高特征提取效率和减少参数量,基于VoVGSCSP及GSConv两种模块实现特征融合网络构建,可保证算法消耗较少内存的同时提高检测精度。

## 1 连铸坯表面缺陷检测轻量化模型构建

为提高算法特征提取效率和减少参数量,本文基于GhostNet设计主干特征提取网络并基于VoVGSCSP及GSConv两种模块设计特征融合网络,构建连铸坯表面

缺陷检测轻量化模型。此轻量化模型构建方法不仅能够有效提升算法在实际应用中的运行速度与准确性,还能降低对硬件资源的需求,使得该技术更加适用于工业现场等计算能力受限的环境;此外,通过优化后的网络结构还可以增强模型对于不同类型、尺寸以及复杂背景下缺陷特征的学习能力,进一步提高了检测系统的鲁棒性和适应性。

### 1.1 连铸坯表面缺陷检测整体网络结构模型

通过预试验,连铸坯表面缺陷检测深度学习模型轻量化的整体网络结构如图1所示。本文基于GhostConv和GhostConv3两种模块设计主干特征提取网络,网络末端通过嵌入快速空间金字塔池化模块以提高检测性能;并基于VoVGSCSP和GSConv两种模块设计特征融合网络,降低模型内存消耗和计算时间;通过头部结构延续一阶段目标检测算法思想,直接对目标信息进行回归。

### 1.2 基于 GhostNet 主干特征提取网络

GhostNet<sup>[13]</sup>作为一种轻量级神经网络架构,通过Ghost模块和GhostBottleNeck实现高效特征提取。其能够在减小模型大小和计算成本的同时保持较好的检测性能。本文基于GhostNet设计Ghost卷积模块,以减少模型计算负担,提高模型计算效率。

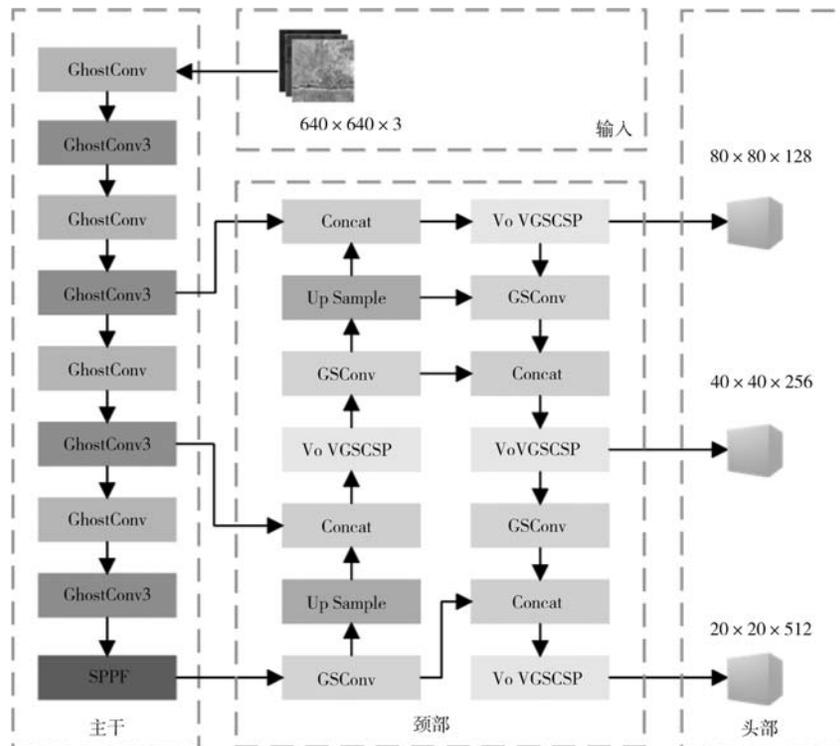


图1 连铸坯表面缺陷检测深度学习轻量化整体网络结构

Fig. 1 Deep learning lightweight overall network structure for the surface defect detection of the continuous casting slab

其中, Ghost卷积模块由以下几部分组成: 首先, 应用 $1 \times 1$ 的普通卷积对图像进行运算, 以达到降低特征图通道数的目的; 其次, 基于前序卷积结果应用分组线性卷积操作, 从而得到新的冗余特征图; 最后, 将第一部分生成的具有特定通道数的特征图与前序的冗余特征图组合, 生成具有特定通道数的特征图(其通道数为 $n$ ), 其实现结构图如图2<sup>[14]</sup>。

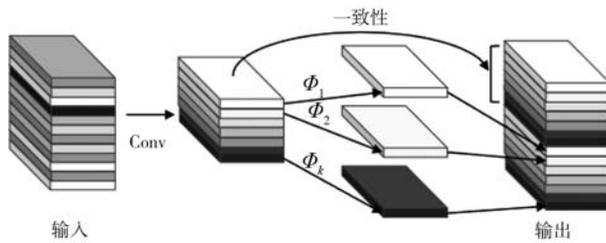


图2 Ghost模块原理图

Fig. 2 Schematic diagram of the ghost module block

为生成通道数为 $n$ 的特定特征图, Ghost模块需先得到一定数量, 如 $s$ 个冗余的特征图, 其计算公式为:

$$y_{i,j} = \varphi_{i,j}(y'_i), \quad \forall i=1, \dots, m, j=1, \dots, s(1)$$

式中:  $y'_i$ 为第 $i$ 个需输入的卷积特征图,  $\varphi_{i,j}$ 为第 $j$ 个线性卷积, 以得到第 $j$ 个冗余卷积特征图 $y_{i,j}$ , 其中得到的第 $s$ 个线性卷积操作 $\varphi_{i,j}$ 为恒等映射。

同时, 对获得的尺寸大小为 $c \times h \times w$ 的卷积特征图, 其对应输出尺寸为 $n \times h' \times w'$ 的卷积特征图, 并使用Ghost操作, 由于包含了恒等映射卷积映射和如 $n(s-1)$ 个线性卷积操作, 则 $F_{\text{Ghost}}$ 的计算表示式如下:

$$F_{\text{Ghost}} = \frac{n}{s} \cdot h \cdot w \cdot c \cdot k \cdot k + (s-1) \cdot \frac{n}{s} \cdot h' \cdot w' \cdot d \cdot d \quad (2)$$

式中:  $w$ 为输入的卷积特征图宽度,  $h$ 为输入的卷积特征图高度,  $c$ 为输入的卷积特征图通道数,  $w'$ 为输出的卷积特征图宽度,  $h'$ 为输出的卷积特征图高度,  $n$ 为输出的卷积特征图通道数,  $k \times k$ 为恒等映射卷积核大小,  $d \times d$ 为线性卷积操作卷积核大小。

对于普通卷积, 其卷积核尺寸为 $n \times c \times k \times k$ , 则 $F_{\text{Conv}}$ 的计算表示式如下:

$$F_{\text{Conv}} = n \cdot h \cdot w \cdot c \cdot k \cdot k \quad (3)$$

根据公式(2)和公式(3), 普通卷积与Ghost卷积的比值 $r_s$ 可表示如下:

$$r_s = \frac{c \cdot k \cdot k}{\frac{1}{s} \cdot c \cdot k \cdot k + \frac{s-1}{s} \cdot d \cdot d} \approx \frac{s \cdot c}{s+c-1} \approx s \quad (4)$$

根据公式(4), 当输出卷积特征图大小相等的情况下, Ghost在特征提取时所需的参数量为Conv卷积的 $1/s$ 倍。即在得到相同数量特征图的情况下, Ghost卷积相比于传统卷积所需的参数更少, 从而减少了计算资源的需求<sup>[15]</sup>。

而GhostBottleneck模块由GhostConv构成。该模块借鉴了ResNet模块的残差结构模型, 并由GhostConv替代一般Conv卷积, 以减少模型的计算量和降低模型内存需求, 如图3。当步长为1时, GhostBottleneck模块由两个GhostConv模块构成, 其中前者模块用于扩展网络通道数, 后者模块用于降低通道数, 以匹配残差连接。当步长为2时, 该GhostBottleneck模块应用步长为2的DWConv连接2个GhostConv卷积。还有, GhostBottleneck模块提取MobileNetv2的设计优点, 只在第一个GhostConv模块后面应用ReLU, 而其他模块后面都不应用激活函数, 同时在每个模块后面均用BN, 从而加快模型的收敛速度并提高收敛效果。

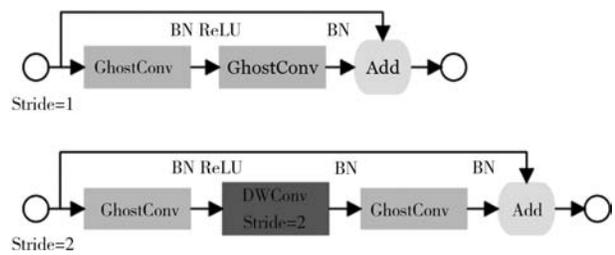


图3 GhostBottleneck结构

Fig. 3 GhostBottleneck structure

整体网络结构中, GhostConv3是基于GhostBottleneck模块构建的结构, 将GhostBottleneck与三个传统的Conv模块组合, 构成的GhostConv3模块如图4所示。

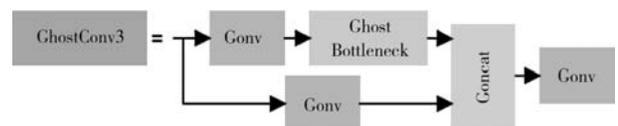


图4 GhostConv3结构

Fig. 4 GhostConv3 structure

### 1.3 基于GSConv降低主干网络模型复杂度

为了降低本算法的运行成本, 借鉴深度可分离卷积层(Depth-wise Separable Convolution, DSC)的优点以降低参数量和计算量。但使用大量DSC的轻量级模型难以获取较好的准确性。因此, 引入新型卷积操作GSConv<sup>[16]</sup>, 以代替模型中的普通卷积层, 从而降低模型复杂度并获取较好的准确性。

其中, GSConv结构模块如图5所示, 其主要包含标准卷积层(Conv)和深度可分离卷积层(DWConv)。首先, 应用标准卷积层对GSConv实施卷积下采样, 得到SC(Standard Convolution)特征图。随后, 对SC进行DWConv卷积, 从而生成DSC特征图。将这两个卷积结果进行拼接, 再经过通道随

机混合操作CSO (Channel Shuffle Operation), 通过Shuffle操作将Conv生成的信息融合到DSC生成的信息中, 以增强信息的互通性。

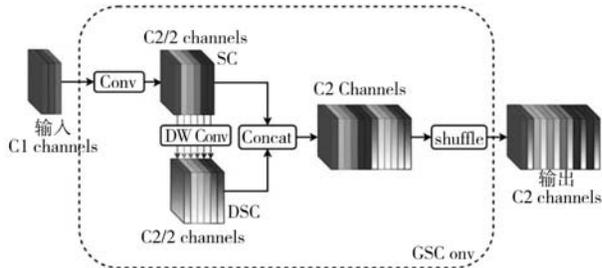


图5 GSConv网络结构

Fig. 5 GSConv network structure

在此模型的所有阶段均使用GSConv, 以使模型的网络层较深, 增加推理时间。因此, 本文仅在特征融合网络部分使用GSConv, 以减轻网络的复杂度, 并提高检测精度<sup>[17]</sup>。

#### 1.4 基于VoVGSCSP构建特征融合网络模型

本文在GSConv的基础上继续引入GSBottleneck与VoVGSCSP两种模块, 构建基于VoVGSCSP与GSConv的特征融合网络模型, 以增强模型的特征融合, 从而提高模型的准确性。

如图6所示, GSBottleneck模块包含两条支路: 一条由两个GSConv模块组成, 另一条由一个Conv模块组成。在卷积过程中, 两条支路能够避免权重共享操作, 并通过通道数的拆分实现不同路径之间的信息传递。有助于减少计算量并提高信息传递的准确性, 从而改善网络整体的性能。

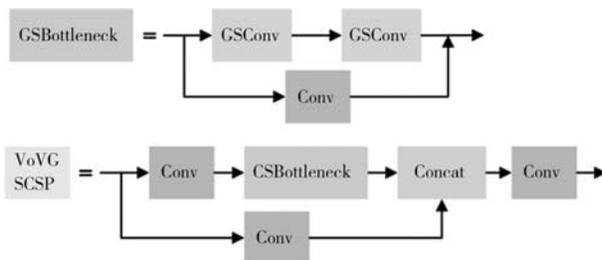


图6 GSBottleneck及VoVGSCSP网络结构

Fig. 6 Network structures of GSBottleneck and VoVGSCSP

VoVGSCSP是一种基于单次聚合的跨阶段局部网络结构, 由GSBottleneck和Conv两种模块组成, 如图6所示。当输入数据特征图进入VoVGSCSP模块时, 一部分通道数通过Conv模块卷积, 然后通过GSBottleneck模块提取特征信息; 另一部分通道数则直接通过Conv模块卷积。这两部分的输出在拼接操作后, 再经过一个Conv模块卷积输出。VoVGSCSP模块综合了GSConv

和GSBottleneck两种模块的优势, 在保持模型轻量化的同时, 能够有效提高检测准确性, 并避免梯度消失问题。

## 2 试验结果与分析

### 2.1 连铸坯表面缺陷分类

由于连铸工艺和其他条件的不同, 连铸坯表面缺陷的种类多样, 且不同行业的标准也不尽相同。本文基于国内外文献及相关标准<sup>[18-19]</sup>, 将缺陷主要划分为裂纹 (crazing)、气孔 (pore) 和夹杂 (inclusion) 三类。

其中, 裂纹是连铸坯最常见的缺陷之一, 通常由于冷却过程中温度梯度不均或应力集中导致。裂纹的形态多种多样, 可能呈现为纵向、横向或网状分布。裂纹的存在不仅会降低材料的力学性能, 还会在后续加工过程中扩展, 导致成品率下降。气孔缺陷是由于铸造过程中气体未能及时逸出, 凝固时被困在金属内部或表面形成。气孔的形状通常为圆形或椭圆形, 内部为空洞。气孔不仅影响材料的致密性, 还可能成为应力集中点, 降低材料的强度和疲劳寿命。夹杂物是指在金属基体中存在的非金属相, 通常由氧化物、硫化物或其他杂质形成。夹杂物的形态复杂多样。夹杂物的存在会导致材料的力学性能下降, 特别是韧性和塑性, 同时也会影响后续加工性能和表面质量。

### 2.2 连铸坯表面缺陷数据集构建

本试验应用中国重型机械研究院股份公司提供的连铸坯表面缺陷数据集, 以应用此实际连铸产线采集的数据集进行试验验证。共采集连铸冷坯表面缺陷图片1 200幅, 其中, 裂纹类图片共460幅, 气孔类图片400幅, 夹杂类图片340幅, 如图7所示为各类表面缺陷的示例。在进行试验数据划分时, 按照一定比例将实验数据集划分为训练数据集和测试数据集两类, 则得到训练数据集900幅和测试数据集300幅, 最后使用Labelling标注软件实现试验数据集的标注。

### 2.3 试验训练环境及参数设置

在进行实验平台构建时, 构建了以Intel (R) Core (TM) i5-12490F @3.0/4.6 GHz为处理器、以NVIDIA GeForce RTX 3060 (12GB) 显卡为加速硬件。而软件支撑平台应用CUDA 11.3为加速软件, Windows 10为操作系统支撑平台。以Pytorch 1.10.0版本为网络模型框架, 编程语言应用Python 3.8。

为了避免模型陷入局部最优并提升检测性能, 训练采用随机梯度下降优化器, 设置模型的初始学习率为0.01, 同时设置模型的动量为0.937, 设置试验模型的批量为8。基于构建的硬件和软件平台系统, 应用采

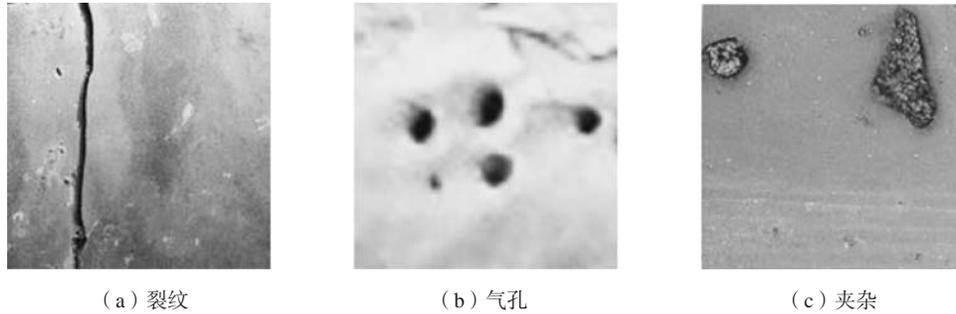


图7 各类连铸坯表面缺陷图像数据集示例

Fig. 7 Examples of various types of surface defect images of the continuous casting slabs dataset

集的各类数据集，通过多轮次预训练和验证，设置试验的总轮次为200。同时采用基于Mosaic的数据增强方法，对输入图像进行随机缩放、裁剪和拼接等数据处理，以提高数据集的多样性及模型的鲁棒性。

## 2.4 试验结果分析与对比

### 2.4.1 评价指标

本文主要采用精确率 $P$ 、召回率 $R$ 和帧率 $R_f$ 作为评价指标。

根据样本的判断结果属性，将样本分为四类：正样本预测为正（True Positives,  $TP$ ）、负样本被预测为负（True Negatives,  $TN$ ）、负样本被预测为正（False Positives,  $FP$ ）、正样本被预测为负（False Negatives,  $FN$ ）。

其中，精确率（Precision,  $P$ ）指正确预测的正样本占所有被预测为正样本比例：

$$P = \frac{TP}{TP+FP} \quad (5)$$

召回率（Recall,  $R$ ）指正确预测的正样本占所有实际为正样本比例：

$$R = \frac{TP}{TP+FN} \quad (6)$$

为使评价更全面，构建平均精确度 $AP$ （Average Precision）和平均精度均值 $mAP$ （mean Average Precision）两个指标分别来评价算法针对单个目标和所有目标的检测精度。其中， $AP$ 指以 $P$ 为纵坐标、 $R$ 为横坐标的二维 $P$ - $R$ 曲线下的面积。 $mAP$ 则指各类平均精确度取均值，其计算公式如下：

$$mAP = \frac{\sum AP}{N} \quad (7)$$

式中： $N$ 为类别数；一般使用重叠度（Intersection Over Union, IOU）阈值为0.5来判断目标检测的正确性， $mAP$ 记作 $mAP@0.5$ 。

$R_f$ 指每秒传输的帧数，在目标检测中表示模型每秒处理的图像数量，用于衡量检测速度。

同时，为描述轻量化方法的复杂度优势，引入模型权重大小（Model Weight Size,  $MWS$ ），用来描述算法训练后权重的大小；参数量（Parameters,  $Param$ ），用来描述模型所包含参数的数量，通常情况下，参数量越多模型越复杂，检测精度越高，但所需的计算资源也越高；浮点运算数（FLOPs），用来描述算法复杂度的衡量指标，以衡量模型在推理阶段的计算量。

### 2.4.2 轻量化模型主干网络对比试验

为验证本文轻量化主干网络优势，设计了面向连铸坯表面缺陷数据集的主干网络对比实验。在所有算法超参数相同的条件下，将目前主流的轻量化主干网络，即轻量化Backbone如：GhostNetBackbone、MobileNetV3Backbone、RepVGGBackbone及ShuffleNetV2Backbone，分别嵌入本文轻量化特征融合网络中进行对比，试验结果如表1所示。

由表1可知，在本文轻量化特征融合网络的搭建条件下，基于GhostConv轻量化主干网络的平均精度均值为0.773，在对比试验中数值最高，检测精度优于其他三种轻量化主干网络，且各类缺陷的平均精确度表现也较为优秀。同时，其帧率为138.89帧/s，检测速度仅次于基于ShuffleNetV2的轻量化主干网络。综合来看，面向连铸坯表面缺陷数据时，基于GhostNet的轻量化主干网络检测性能最优，能够在保持一定检测速度的同时，实现连铸坯表面缺陷关键特征信息的高效提取。此外，试验结果也验证了模型权重大小、模型参数量和浮点运算次数三个评价指标仅是检测速度的影响因素之一。

为验证轻量化连铸坯表面缺陷检测方法的检测性能优势，将基于本文轻量化主干网络分别嵌入YOLOv5算法的特征融合网络（YOLONeck）和轻量化特征融合网络SlimNeck，面向连铸坯表面缺陷数据集使用开源算法生成混淆矩阵，结果如图8所示。

由图8（a）和（b）可知，相比YOLONeck，本文设计的轻量化特征融合网络SlimNeck在面向连铸坯

表1 基于轻量化特征融合网络的主干网络对比试验  
Tab. 1 Comparison experiments of the backbone networks based on the lightweight feature fusion networks

算法	$mAP@0.5$	$AP@0.5$			$R_f/(\text{帧}\cdot\text{s}^{-1})$	$MWS/\text{MB}$	$Params/10^6$	$FLOPs/\text{G}$
		裂纹	气孔	夹杂				
GhostNetBackbone+SlimNeck	0.773	0.873	0.573	0.872	138.89	8.3	3.91	7.4
MobileNetV3Backbone+SlimNeck	0.759	0.865	0.567	0.846	105.26	5.4	2.45	3.4
RepVGGBackbone+SlimNeck	0.757	0.832	0.565	0.876	135.14	3.6	1.59	5.4
ShuffleNetV2Backbone+SlimNeck	0.693	0.767	0.519	0.791	149.25	1.5	0.56	1.1

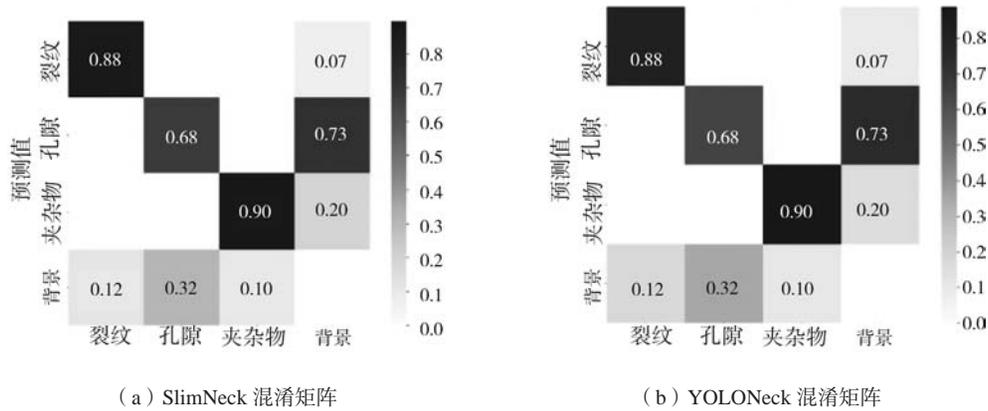


图8 面向连铸坯表面缺陷数据集的混淆矩阵

Fig. 8 Confusion matrix for the surface defect dataset of the continuous casting slabs

表面缺陷数据集时，缺陷检测成功概率明显提高。其中，裂纹、气孔和夹杂三类缺陷检测成功概率分别为0.88、0.68和0.90，且三类缺陷检测成功概率分别提高7.3%、4.6%和1.1%。裂纹、气孔和夹杂三类缺陷检测为背景（漏检）概率分别为0.12、0.32和0.10。背景检测为缺陷（误检）概率，裂纹、气孔和夹杂三类缺陷分别为0.07、0.73和0.20。混淆矩阵结果表明，相较YOLONeck，本文设计的轻量化特征融合网络在面向连铸坯表面缺陷检测任务时检测性能更优，能够一定程度上避免漏检和误检情况的出现。

#### 2.4.3 轻量化模型不同数据集对比试验

##### (1) 连铸坯表面缺陷数据集对比试验

为验证本文轻量化算法优势，此处面向连铸坯表面

缺陷数据集设立对比试验，将其与目前常见轻量化算法进行比较，为确保试验具有可比性，在所有模型超参数相同的条件下进行对比试验。试验结果如表2所示。

由表2可知，YOLOv4-tiny的帧率为130.29帧/s，检测速度表现优异，但平均精度均值为0.626，检测精度较低；YOLOv5s和YOLOv7-tiny两种算法性能相似，其中帧率分别为123.46帧/s和116.28帧/s，平均精度均值分别为0.746和0.756；YOLOX-tiny的平均精度均值为0.765，检测精度表现接近本文轻量化算法，但帧率仅为98.04帧/s，检测速度较低；本文轻量化算法的平均精度均值为0.773，检测精度最高，各类缺陷平均精度表现优秀，且帧率为138.89帧/s，检测速度最高，模型权重大小、模型参数量和浮点运算次数分别为8.3 MB、 $3.91 \times 10^6$ 和7.4 G。试验结果表明，本文轻量化算法综

表2 面向连铸坯表面缺陷数据集的对比试验结果  
Tab. 2 Comparative experimental results for the surface defect dataset of the continuous casting slabs

算法	$mAP@0.5$	$AP@0.5$			$R_f/(\text{帧}\cdot\text{s}^{-1})$	$MWS/\text{MB}$	$Params/10^6$	$FLOPs/\text{G}$
		裂纹	气孔	夹杂				
YOLOv4-tiny	0.626	0.743	0.389	0.745	130.29	22.5	5.89	6.8
YOLOv5s	0.746	0.859	0.476	0.903	123.46	14.5	7.02	15.8
YOLOX-tiny	0.765	0.867	0.564	0.865	98.04	19.4	5.03	6.4
YOLOv7-tiny	0.756	0.882	0.533	0.855	116.28	11.7	6.02	13.1
本文轻量化算法	0.773	0.873	0.573	0.872	138.89	8.3	3.91	7.4

合性能最优。

#### (2) 热轧带钢表面缺陷数据集对比试验

为了验证构建的轻量化模型算法在连铸工艺生产的连铸坯表面缺陷检测中的泛化性等模型特性, 首先

应用热轧带钢表面缺陷数据集进行试验。将其与目前常见轻量化算法进行比较, 试验结果如表3所示。

由表3和表4可知, 基于此数据集试验结果, 本文轻量化算法的平均精度均值为0.770, 检测精度最高,

表3 热轧带钢表面缺陷数据集对比试验结果 (精度)

Tab. 3 Comparative experimental results on the surface defect dataset of the hot-rolled strip (accuracy)

算法	mAP@0.5	AP@0.5					
		裂纹	夹杂	斑块	表面点蚀	轧制氧化皮	划痕
YOLOv4-tiny	0.650	0.302	0.767	0.900	0.622	0.451	0.859
YOLOv5s	0.749	0.487	0.789	0.948	0.831	0.563	0.878
YOLOX-tiny	0.765	0.396	0.823	0.970	0.829	0.711	0.863
YOLOv7-tiny	0.770	0.345	0.963	0.803	0.959	0.886	0.665
本文轻量化算法	0.770	0.513	0.759	0.963	0.823	0.687	0.876

表4 热轧带钢表面缺陷数据集对比试验结果 (速度)

Tab. 4 Comparative experimental results on the surface defect dataset of the hot-rolled strip (speed)

算法	$R_p/(\text{帧} \cdot \text{s}^{-1})$	MWS/MB	Params/ $10^6$	FLOPs/G
YOLOv4-tiny	134.96	22.5	5.89	6.8
YOLOv5s	138.89	14.5	7.02	15.8
YOLOX-tiny	91.03	19.4	5.03	6.4
YOLOv7-tiny	86.21	11.7	6.02	13.1
本文轻量化算法	151.52	8.3	3.91	7.4

各类缺陷平均精度表现优秀, 且帧率为151.52帧/s, 检测速度最高。模型权重大小、模型参数数量和浮点运算次数表现优异, 模型复杂度较低, 其分别为8.3 MB、 $3.91 \times 10^6$ 和7.4 G。试验结果表明, 面向热轧带钢表面缺陷数据集时, 本文构建的轻量化模型算法在综合性能上表现最佳, 表明该模型算法在连铸坯表面缺陷检测领域具有良好的泛化能力和普适性, 可为计算资源受限条件下的算法应用部署问题提供一定解决思路。

本文通过上面对比试验验证了所提出的轻量化算法在连铸坯表面缺陷检测中的有效性和优越性。首先, 主干网络对比试验表明, 基于GhostNet的轻量化主干网络在检测精度上表现最佳, 平均精度均值达到0.773, 同时保持了较高的检测速度(138.89帧/s); 其次, 针对连铸坯表面缺陷数据集的实验结果显示, 本文轻量化算法在综合性能上优于其他常见轻量化算法, 如YOLOv4-tiny、YOLOv5s、YOLOv7-tiny和

YOLOX-tiny, 具有更高的检测精度(平均精度均值为0.773)和较快的检测速度(138.89帧/s); 最后, 面向热轧带钢表面缺陷数据集的试验进一步证明了本文算法的普适性和通用性, 其平均精度均值为0.770, 检测速度为151.52帧/s, 显示出优异的综合性能。这些试验结果表明, 本文提出的轻量化算法在金属冶金表面缺陷检测领域具有显著的优势和应用潜力。

### 3 结束语

本研究提出的主干特征提取与融合的连铸坯表面缺陷检测轻量化模型方法中, 试验采集的连铸坯表面质量缺陷应用数据集主要来自连铸工程实际应用的生 产线, 并结合国内外的表面缺陷分类文献和行业标准缺陷分类文献, 未经过相应的数据增强以及亮度调节等数据预处理, 以更贴近实际工程应用, 确保结果具有较高的可靠性。根据提出的基于主干特征提取与融合的轻量化模型方法在实际工程应用的连铸坯表面缺陷检测的应用效果, 与现有的常规的智能检测算法及对比试验中的其他算法相比, 该方法能够更好地平衡检测的精确率和检测速度之间的关系。由于高温和氧化皮等因素, 连铸坯在线检测面临一定的挑战。后续的工作将通过迁移学习来处理不同环境下的连铸坯表面缺陷, 以提高模型的鲁棒性和适用性, 同时研究相关硬件系统, 以实现复杂环境下的在线缺陷检测。

#### 参考文献:

- [1] 朱苗勇, 姜文涛, 王卫领. 炼钢与连铸过程数值模拟研究进展 [J]. 金属学报, 2018, 54 (2): 131-150.
- [2] 黄军, 王宝峰, 张雪元, 等. 高温铸坯表面缺陷在线检测技术开发及应用 [J]. 连铸, 2022 (6): 61-67.
- [3] 赵立明. 基于激光扫描成像与异源CCD融合的连铸坯表面缺陷检测方法研究 [D]. 重庆: 重庆大学, 2014.
- [4] 邵长禄, 姜浩. 超声技术对连铸圆坯内部质量分级的研究 [J]. 无损探伤, 2019, 43 (6): 32-33.

- [5] 张雪元. 基于激光扫描方法的高温铸坯表面缺陷检测研究与应用 [D]. 包头: 内蒙古科技大学, 2021.
- [6] 徐科, 杨朝霖, 周鹏. 热轧带钢表面缺陷在线检测的方法与工业应用 [J]. 机械工程学报, 2009, 45 (4): 111-114, 124.
- [7] 胡嘉成, 王向阳, 刘晗. 基于深度学习的连铸坯表面缺陷检测 [J]. 上海大学学报 (自然科学版), 2019, 25 (4): 445-452.
- [8] 王睿泽, 孙铁强, 宋超, 等. 基于改进YOLOv3算法的连铸坯表面缺陷检测 [J]. 组合机床与自动化加工技术, 2023, 587 (1): 75-78.
- [9] 成彬, 王井浩, 何博, 等. 基于FN-YOLOv5的连铸坯表面缺陷检测方法 [J/OL]. 材料科学与工艺, 1-10[2024-09-20].
- [10] XIE W, SUN X, MA W. A light weight multi-scale feature fusion steel surface defect detection model based on YOLOv8 [J]. Measurement Science and Technology, 2024: 055017.
- [11] ZHANG L, FU Z, GUO H, et al. Multiscale local and global feature fusion for the detection of steel surface defects [J]. Electronics, 2023, 12 (14): 3090.
- [12] LIANSHAN S, JINGXUE W, HANCHAO D, et al. MSFF: A multi-scale feature fusion network for surface defect detection of aluminum profiles [J]. IEICE Transactions on Information and Systems, 2022, E105. D (9): 1652-1655.
- [13] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations [C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1580-1589.
- [14] 孙月平, 袁必康, 詹婷婷, 等. 基于改进YOLOv5s的河蟹与饵料检测方法 [J]. 农业工程学报, 2023, 39 (22): 178-187.
- [15] 杨永波, 李栋, 房建东, 等. 面向嵌入式端的轻量级交通信号灯检测算法 [J]. 计算机工程与应用, 2024, 60 (13): 361-368.
- [16] LI H, LI J, WEI H, et al. Slim-neck by GSConv: a better design paradigm of detector architectures for autonomous vehicles. arXiv 2022 [J]. arXiv preprint arXiv: 2206.02424.
- [17] 高英, 吴玉虹. 基于YOLOv5s的轻量化行人检测算法 [J]. 现代电子技术, 2023, 46 (22): 151-158.
- [18] STALHEIM G D, 王厚昕. 连铸板坯和方坯表面缺陷的分析与判定 [J]. 世界钢铁, 2010, 10 (4): 23-28.
- [19] 全国钢标准化技术委员会. YB/T 4003—2016连铸钢板坯低倍组织缺陷评级图 [S]. 北京: 中国标准出版社, 2017.

---

## Research on the Lightweight Model for Surface Defect Detection of Continuous Casting Slab Based on Main Feature Extraction and Fusion

CHENG Bin, JING Chang-xin, WANG Jing-hao, LU Bo

(School of Mechanical and Electrical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China)

### Abstract:

To address the challenge of limited computational resources, a method based on lightweight deep learning models for detecting surface defects on continuous casting billets has been proposed in this paper. This approach aims to achieve efficient and rapid defect detection. Firstly, a backbone feature extraction network based on GhostNet was designed, which significantly improved the efficiency of feature extraction while greatly reducing the number of network parameters. Subsequently, a feature fusion network incorporating VoVGSCSP and GSConv modules was constructed, optimizing the feature fusion strategy to further reduce the model's memory usage and computational overhead. Finally, comparative experiments were conducted using datasets of continuous casting billet surface defects and hot-rolled strip steel surface defects. It was demonstrated by the experimental results that the proposed method achieved excellent performance across multiple metrics: the average detection accuracy reached 0.773, the detection speed was increased to 138.89 frames per second; the model size was 8.3 MB with  $3.91 \times 10^6$  parameters, and the floating-point operations (FLOPs) were reduced to 7.4 G. Compared to current mainstream models, this method exhibited a significant advantage in terms of complexity reduction, achieving a good balance between detection speed and accuracy. It provides a highly valuable new approach and technical implementation scheme for the research of deep learning algorithms for the surface defect detection on the continuous casting billets under resource-constrained conditions.

### Key words:

continuous casting; surface defects; intelligent detection; lightweight; feature extraction and fusion