

钢材缺陷检测算法描述文档

1 引言

钢材表面缺陷检测在钢铁生产中是确保质量的关键环节，传统的人工检测方式难以满足大规模工业生产的需求。近年来，基于深度学习的缺陷检测方法因其高效性和准确性受到广泛关注。然而，现有的深度学习模型如 U-Net 虽具备较好的分割性能，但由于参数量大、计算复杂度高，限制了其在实际生产环境中的应用。为此，本文提出了一种基于轻量级 U-Net 的改进模型，用于钢材表面缺陷检测。

2 系统框架

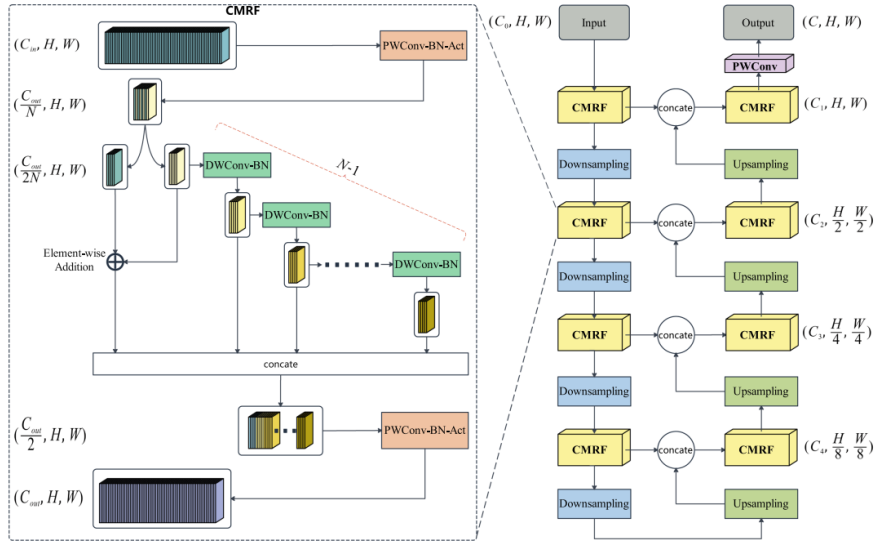


图 1: 钢材缺陷检测算法框图

3 算法描述

我们的方法基于轻量化卷积神经网络结构，通过一系列数据增强和自适应优化技术提升模型的泛化能力。算法流程如下：

1. 数据加载与预处理。

2. 数据增强以提高模型在训练过程中的鲁棒性。
3. 使用自适应权重更新策略训练模型。
4. 在验证集上计算 IoU、精度等指标进行性能评估。
5. 将本模型与基线模型进行对比分析。

4 模型设计思路

本模型设计采用轻量化架构，通过不同层次的特征融合和优化策略提升模型在实际工业场景中的应用效果。下文从数据增强、模型架构、损失函数、优化方法四个方面进行详细说明。

4.1 数据增强技术

我们在数据加载阶段应用多种数据增强技术，使模型在不同数据条件下具有更强的适应性。以下是主要数据增强操作：

- **灰度转换**：将输入图像转为灰度图 $I_{gray} = 0.299I_R + 0.587I_G + 0.114I_B$ ，以降低计算复杂度。
- **随机翻转和旋转**：为增强数据的多样性，我们应用了随机水平翻转与 90 度旋转。
- **噪声添加**：高斯噪声 $n \sim \mathcal{N}(0, \sigma^2)$ 使模型更具鲁棒性。

4.2 模型结构设计

本模型的架构主要包含多接收域特征提取模块 (CMRF)，该模块结合了深度卷积和点卷积，能够高效提取多层次特征。模型包含编码器和解码器两部分，通过多接收域的设计增强模型特征提取的多样性。

4.2.1 多接收域特征模块 (CMRF)

在本模型中，我们设计了多接收域特征模块，用于从多尺度提取特征信息。给定输入特征图 $X_{in} \in \mathbb{R}^{C_{in} \times H \times W}$ ，CMRF 模块的工作流程如下：

1. **点卷积**：通过点卷积 (Pointwise Convolution, PWConv) 将输入特征通道调整到目标通道数：

$$X' = \sigma(W_{pw} * X_{in} + b_{pw})$$

其中， σ 表示激活函数（本文使用 GELU 激活）。

2. **深度卷积**：使用深度卷积 (Depthwise Convolution, DWConv) 对特征进行空间信息提取。对于每个通道 c ，计算如下：

$$X''_c = W_{dw,c} * X'_c$$

其中， $*$ 表示卷积运算。

3. **级联策略**：将不同卷积核大小的特征拼接，以提高特征的多样性。

4.2.2 编码器和解码器

模型的编码器与解码器采用对称设计，其中编码器主要进行下采样操作，通过 CMRF 模块逐步提取高级特征；解码器通过上采样与特征融合还原空间信息。给定输入特征图 $Y_{in} \in \mathbb{R}^{C_{in} \times H \times W}$ ，编码器和解码器的运算过程如下：

$$Y_{enc} = \text{DWConv}(Y_{in}) + \text{PWConv}(Y_{in}) \quad (1)$$

$$Y_{dec} = \text{Upsample}(Y_{enc}) + \text{Concat}([Y_{enc}, Y_{skip}]) \quad (2)$$

其中， Y_{skip} 表示跳跃连接部分的特征。

5 损失函数与优化策略

在损失函数方面，采用了标准交叉熵损失与 Lovasz Softmax 损失的组合，以增强对小目标的检测能力。损失函数定义如下：

$$L = \alpha \cdot L_{\text{CE}} + \beta \cdot L_{\text{Lovasz}} \quad (3)$$

其中， L_{CE} 为交叉熵损失， L_{Lovasz} 为 Lovasz 损失， α 和 β 为权重系数。

此外，我们使用自适应学习率调整策略 (ReduceLROnPlateau)，在模型性能不提升时降低学习率，公式如下：

$$\eta_{t+1} = \eta_t \times \gamma$$

其中， η 为学习率， γ 为学习率衰减系数。

6 实验结果与分析

表 1 展示了本模型与基线模型在性能、参数量和 FPS 上的对比。

模型	参数量 (M)	类别 1 IoU	类别 2 IoU	类别 3 IoU	mIoU (%)	mPA (%)	FPS
基线模型	17.262	0.6790	0.8369	0.7840	76.67	89.89	32
我们的模型	0.4813	0.7869	0.9028	0.8480	84.59	93.19	143.42

表 1: 基线模型与我们模型的性能、参数量及 FPS 对比

如表 1 所示，我们的模型在 mIoU 和 mPA 指标上分别达到 84.59% 和 93.19%，显著优于基线模型。此外，该模型的参数量仅为 0.4813M，相较基线模型的 17.262M 减少了 97%，计算开销大幅降低。FPS 方面，我们的模型达到了 143.42 FPS，表明其在实时检测中的潜力。