

# 基于改进 YOLOv11 的短波信号识别方法研究

李佳峰<sup>1,2</sup>, 张海瑛<sup>1</sup>, 王晓亚<sup>1</sup>, 李翊嘉<sup>1</sup>

(1. 中国电子科技集团公司 第五十四研究所, 石家庄 050081;

2. 河北省电磁频谱认知与管控重点实验室, 石家庄 050081)

**摘要:** 针对短波信号侦察中信号识别效果不佳的问题, 提出了一种基于信号时频图像和 WT-YOLO 的短波信号识别方法; 该方法将短波信号通过短时傅里叶 STFT 变换变成信号时频图形式; 针对信号低频语义信息与高频细节互相干扰问题, 在模型中引入基于小波变换上采样 WFU 模块来提高模型的特征融合能力; 设计了三重感受野 TRF 模块, 解决单一感受野无法多特征提取的问题; 引入了 PIoUv2 模块提高模型定位精度, 从而提高模型检测识别精度; 实验结果表明, WT-YOLO 比原有网络模型有更高的识别准确率, 达到 96.4%, 识别错误率相对下降 55.6%。

**关键词:** 短波信号识别; 时频图; 小波变换; 三重感受野; WT-YOLO

## Research on Shortwave Signal Recognition Method Based on Improved YOLOv11

LI Jiafeng<sup>1,2</sup>, ZHANG Haiying<sup>1</sup>, WANG Xiaoya<sup>1</sup>, LI Yijia<sup>1</sup>

(1. The 54th Research Institution of CETC, Shijiazhuang, Hebei 050081, China;

2. Hebei Province Key Laboratory of Electromagnetic Spectrum Cognition and Control, Shijiazhuang, Hebei 050081, China)

**Abstract:** Aiming at the problem of poor signal recognition effect in shortwave signal reconnaissance, a shortwave signal recognition method based on signal time-frequency image and WT-YOLO is proposed. This method transforms shortwave signals into signal time-frequency image form through short-time Fourier transform (STFT). To address the issue of mutual interference between low-frequency semantic information and high-frequency details of the signal, a wavelet transform up-sampling (WFU) module is introduced into the model to enhance the feature fusion ability of the model. A triple receptive field (TRF) module is designed to solve the problem that a single receptive field cannot extract multiple features. The PIoUv2 module is introduced to improve the model's positioning accuracy, thereby enhancing the detection and recognition accuracy of the model. Experimental results show that WT-YOLO has a higher recognition accuracy of 96.4% compared to the original network model, with a 55.6% relative decrease in recognition error rate.

**Keywords:** shortwave signal recognition; time-frequency diagram; wavelet transform; triple receptive field; WT-YOLO

## 0 引言

短波信号 (3~30 MHz) 凭借其超视距传输、抗毁性强以及全天候工作等特性, 在军事通信、民用航空、应急通信中具有不可替代的战略地位, 尤其是在复杂的电磁环境下, 短波信号可通过电离层反射实现数千公里

内无中继通信<sup>[1-3]</sup>。短波信号侦察识别是指在非合作接收情况下, 从宽带接收信号中检测各类短波信号、估计其带宽、中心频率等信号参数、并实现短波信号的种类识别。短波信号识别技术可用于预测和感知战场态势, 以便掌握战争主动权, 这也是短波侦察对抗领域的重要研究内容<sup>[4]</sup>。因此, 研究更加准确高效的短波信号识别

收稿日期:2024-07-17; 修回日期:2024-09-02。

基金项目:国家自然科学基金(U22B2002)。

作者简介:李佳峰(2001-),男,硕士研究生。

通讯作者:张海瑛(1973-),女,博士,高级工程师。

引用格式:李佳峰,张海瑛,王晓亚,等. 基于改进 YOLOv11 的短波信号识别方法研究[J]. 计算机测量与控制, 2026, 34(1):150

-156,165.

算法变得十分重要。

早期短波信号识别研究主要基于信号特征解析<sup>[5]</sup>和模板匹配<sup>[6]</sup>。文献 [7] 通过设计突发信号模板, 结合滑窗傅里叶变换实现了 3GALE 系统 BW0 突发波形的检测; 文献 [8] 则通过创建特定信号时频图模板, 利用计算与接收信号时频图相关系数完成了多种短波信号鉴别; 文献 [9] 采用频谱特征匹配技术, 通过信号与预设模板的频谱相似度实现了 2GALE 信号的识别; 文献 [10] 采用时域子序列匹配算法识别 STANAG4285 等相移键控 (PSK, phase shift keying) 信号, 并通过频域波峰匹配算法识别 2GALE 等频移键控 (FSK, frequency shift keying) 信号。以上这些方法依赖人工参与特征模板构建和检测阈值设定, 受限性高。

随着深度学习的快速发展, 许多学者逐渐探索将深度学习用于通信信号调制识别领域, 为这一传统领域开辟了全新的技术路径。文献 [11] 首先提出将深度学习用于信号调制识别, 设计卷积神经网络 (CNN, convolutional neural network) 来识别 3 种模拟信号调制和 8 种数字信号调制, 识别效果优于传统算法。文献 [12] 设计了卷积长短期深度神经网络 (CLDNN, convolutional long short-term deep neural networks) 来检测信号是否存在, 但无法确定信号类型以及调制方式。文献 [13] 将从短波信号矢量图和数据序列中提取到的融合特征放入 CNN 网络进行短波信号识别, 实现了 6 种短波信号识别, 但其要求训练样本和测试样本要有完整的信号帧结构, 实际信号侦察中难以获取信号完整帧结构。模型改进是基于深度学习短波识别研究的热点方向。现有模型改进策略主要是在模型中引入注意力机制<sup>[14]</sup>、优化损失函数<sup>[15]</sup>、轻量化机制<sup>[16]</sup>以及加强特征融合机制<sup>[17]</sup>, 模型改进后较基准模型都有所提升, 但还需进一步实验验证其对短波信号识别效果。YOLO 模型同时兼顾识别准确率与效率, 近年来被广泛用于信号识别领域, 朱小俊等人在 YOLOv8 传统模型上删减卷积层数, 实现了模型轻量化。朱政宇等人在 YOLOv8 传统模型上引入可形变卷积核, 提高了模型推理速度, 李润东等人改进 YOLOv3 模型的损失函数, 将平均识别准确率提高 3%, 单阶段检测识别 YOLO 模型凭借较高识别准确率以及较好实时性在信号检测识别领域得到一定应用, 但上述 YOLO 模型并未依据信号协议进行定制化设计, 识别准确率与效率仍有提升的潜力。

在短波信号侦察识别中, 现有固定感受野信号检测识别模型难以兼顾信号多尺度特征, 实际信号侦察识别中, 短波信号会受多径效应影响使得信号时域展宽, 会因多普勒效应引起频率抖动, 会因信号衰弱造成能量衰减, 单一尺度特征提取器对此类非线性形变识别准确率不足。因此, 本文提出一种基于改进 YOLOv11 的短波

信号识别模型, 与改进前 YOLOv11 模型相比, 所提出的改进模型信号虚警概率明显下降。本文具体研究工作如下, 本文分析了短波信号在时频图上呈现出的视觉特征, 构建了用于训练深度学习网络的短波信号时频图数据集, 数据集包含 6 类侦查识别中常见的短波信号, 为了解决信号时频图中高频特征与低频语义混叠的问题, 使用基于小波变换上采样 WFU 模块替代原有最邻近插值上采样, 使模型先将高频特征与低频语义分开处理后再融合, 避免高频与低频特征的混叠, 为解决单一感受野难以兼顾信号在频率轴需要高分辨率特征提取和在时间轴需要长程依赖建模能力的需求, 设计了一种基于不同膨胀率和填充的三重感受野 TRF 模块, 提高模型多特征提取能力, 为提高模型检测识别精度, 引入 PI-oUv2 损失函数, 增强模型定位精度, 从而提高检测识别精度。

## 1 短波信号时频特性分析

在深度学习中, 短波信号数据训练形式有短波 IQ 数据、信号星座图、信号时频图, 其中, 信号时频图能同时提供时域和频域的信息, 这对于非平稳信号非常重要。在短波信号侦察中, 不同的调制信号在不同频段有不同的频率成分, 时频图可以捕捉这种变化, 而单纯的时域或频域分析会丢失这些信息。短波信号协议规定信号的帧结构、调制方式以及载波中心频率, 不同类型的短波信号由于短波信号协议的不同会在时频图上呈现出不同的视觉特征。常用的时频分析方法有短时傅里叶变换<sup>[18]</sup> (STFT, short time fourier transform)、小波变换<sup>[19]</sup> (WT, wavelet transform) 和 Wigner-Ville<sup>[20]</sup> 分布等, 其中, Wigner-Ville 分布虽然理论上具有更优的时频聚集性, 但其存在交叉项干扰, STFT 相较于小波变换和 Wigner-Ville 计算量更低且不存在交叉项干扰, 因此, 本文采用 STFT 来进行信号时频转换。

通过 STFT 变换将短波信号数据由一维数据形式变为二维时频图形式, STFT 定义为:

$$STFT(t, f) = \int_{-\infty}^{\infty} s(\tau) \omega(\tau - t) e^{-j2\pi f\tau} d\tau \quad (1)$$

其中:  $t$  为时间,  $f$  为频率,  $s(t)$  为输入信号,  $\omega(t)$  为窗函数。STFT 原理就是先对信号加窗, 通过窗函数截取信号, 然后进行快速傅里叶变化 (FFT, fast fourier transform) 滑动时间窗口, 最后各个窗函数信号频谱组合起来就是信号时频图。

对短波信号侦察中常见的 LINK11 CLEW<sup>[21]</sup>、CLOVER2000<sup>[22]</sup>、2GALE、STANAG4285、CIS12 和 3GALE 等 6 类短波信号进行时频特征分析。LINK11 CLEW 采用四相相对相移键控调制 (DQPSK, different quadrature phase shift keying) 方式, 其信号帧结构包括 605 Hz 的多普勒校正音频、2 915 Hz 的同步音频以

及频率间隔为 110 Hz 从 935 Hz 到 2 365 Hz 的 14 个信息载波。LINK11 CLEW 信号时频图呈现出 16 条平行谱线, 605 Hz 多普勒校正谱线能量显著高于其他谱线(高 5~7 dB), 形成了“强基准线+密集子带”模式, 其时频图如图 1 (a) 所示。CLOVER2000 采用 8 音调制方式, 协议特征为 8 个单音脉冲, 频率范围为 625 Hz 到 2 375 Hz, 频率间隔为 250 Hz, CLOVER2000 信号时频图显示 8 条等间距谱线, 时域上表现为周期性脉冲, 频域上能量均匀分布, 其时频图如图 1 (b) 所示。2GALE 信号采用 8 FSK 调制方式, 在信号时频图中 2GALE 信号频率跳变轨迹清晰可见, 行程 8 条平行频带, 跳变边缘存在过渡带, 其信号时频图如图 1 (c) 所示。STANAG4285 采用 8 PSK 调制方式, 其时频图是连续频谱, 能量在频带内均匀扩散, 时间包络恒定, 其时频图如图 1 (d) 所示。CIS12 采用 12 音调 PSK 制方式, 其在时频图中呈现 12 条等间距的离散谱线, 能量非均匀分布, 由于 PSK 符号速率低, CIS12 时频图呈现出每条线宽度较窄, 其信号时频图如图 1 (e) 所示。3GALE 信号采用 8 PSK 调制方式, 与 STANAG4285 类似但带宽更宽, 频带中心存在载波能量聚集现象, 其信号时频图如图 1 (f) 所示。

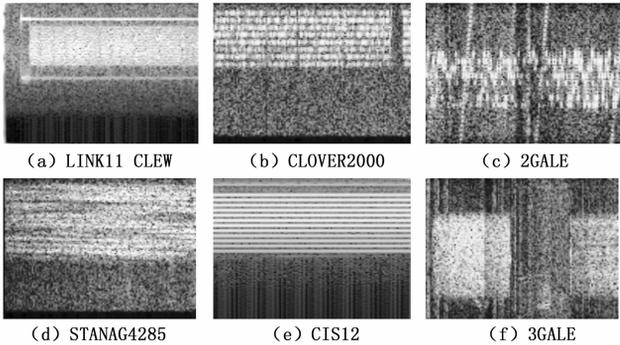


图 1 6 类短波信号时频图

## 2 YOLOv11 模型架构

YOLOv11 是由 Ultralytics 公司于 2024 年发布的模型架构, 其主要由骨干网络 (Backbone)、颈部网络 (Neck) 和检测头 (Head) 三部分组成, 其模型架构如图 2 所示, 其中, 骨干网络部分主要包含 Conv 模块、C3k2 模块、SPPF 模块以及 C2PSA 模块等功能单元, 颈部网络部分主要包含上采样 (Upsample) 模块、C3k2 模块以及 Conv 模块等功能单元。

在 YOLOv11 中骨干网络负责从输入图像中提取多尺度特征, Conv 模块进行初级特征提取, 捕获时频图的局部纹理, 相较于 YOLOv8 模型架构, YOLOv11 模型用 C3k2 单元替换 C2f 单元, 通过动态调整 C3k2 模型参数增强模型特征学习能力, SPPF 模块通过卷积和最大池化操作提高模型整体的感受野, 融合时频图全局

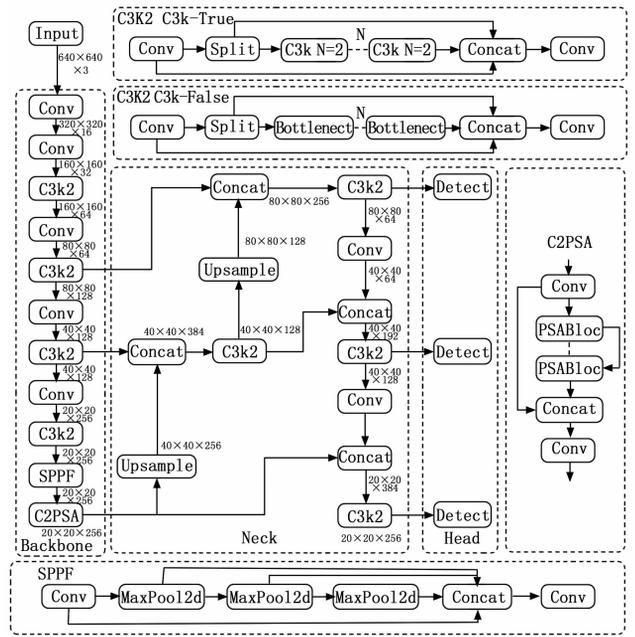


图 2 YOLOv11 模型架构图

上下文信息。颈部网络通过融合不同层级的特征图, 使模型能够捕捉目标特有的视觉属性。

## 3 YOLOv11 改进

### 3.1 小波变换上采样 WFU 模块

在短波信号侦查识别中, 短波信号时频图呈现出多尺度特性, 如 LINK11 CLEW 的 16 条平行谱线、STANAG4285 的连续带宽、2GALE 跳频边缘的过渡带和 CLOVER2000 脉冲的时域过冲, 传统 YOLOv11 采用的最邻近插值方法进行上采样在融合深层次特征时, 低频语义信息与高频细节互相干扰, 导致窄带信号的谱线模糊和跳频信号的边缘失真。为了解决上述问题, 引入基于小波变换的多尺度分解与动态频率融合的方式进行上采样, 替换原模型中 Upsample 模块。

在传统深度学习模型中, 通常需要通过上采样将不同尺度的特征图对齐后进行融合, 然而, 直接融合会导致高频特征和低频特征的混叠, 影响信号时频图细节质量, 现有的方法虽然能传递信息, 但未充分考虑信号时频图不同频率特征的细节, 导致细节恢复不清楚, WFU 通过利用小波变换的多尺度分析能力, 将时频图不同尺度特征分解为高频和低频分量, 分别进行处理和融合, 以避免混叠并增强时频图中信号细节特征。

为了在计算复杂度和保留关键信息之间取得最佳平衡, 在 WFU 模块中选择 Haar 小波作为基函数进行单层分解, WFU 结构如图 3 所示, 首先对来自较大尺度特征  $F_s$  应用小波变换 (WT), 分解为低频分量  $A_{LL}$ 、高频分量  $D_{RR}$ 、高频分量  $V_{RL}$  和高频分量  $H_{LR}$ , 低频分量用于捕获信号整体结构特征, 高频分量用于捕获信号

时频图呈现出的边缘、纹理等细节特征。 $F_{s+1}$  主要包含低频信息, 将其与  $A_{LL}$  串联作为增强后的低频子带, 强化整体结构的连贯性。对三个高频分量  $H_{LR}$ 、 $V_{RL}$  和  $D_{RR}$  通过残差块进一步提取信号时频图细节特征, 抑制噪声并增强边缘响应。

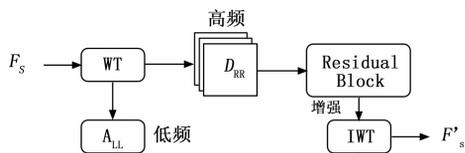


图 3 WFU 结构图

将处理后的低频分量与高频分量通过以 Haar 小波基为基函数的逆小波变换 (IWT) 重构, 生成上采样后的特征, 其变换式如下所示:

$$F'_s = IWT[Concat(A_{LL}^i, F_{s+1}), R(H_{LR}^i, V_{RL}^i, D_{RR}^i)]$$

其中:  $Concat$  表示拼接操作,  $R$  表示残差块操作, 通过这种方式, 基于小波变换的上采样 WFU 在信号识别中, 实现了跨尺度的频率特征分离和融合, 避免了最邻近插值法上采样导致的混叠问题。传统 YOLOv11 模型中通过最邻近插值法上采样, 容易丢失信号时频图中高频细节特征, WFU 利用小波变换的多分辨率特性, 将不同尺度的特征分解为频率分量, 分别处理后再融合, 确保低频结构和高频细节的准确传递。因此, 为了提高短波信号识别精确度和鲁棒性, 本文将 Neck 层中的基于最邻近插值的上采样模块替换为基于小波变换的上采样 WFU 模块, 增强模型对短波信号特征学习能力。

### 3.2 三重感受野 TRF 模块

感受野是深度学习中极其核心且重要的概念, 感受野定义了单个输出单元所能利用的输入信息的空间范围。由于短波信号在时频图中表现为不同尺度特征 (如瞬态脉冲、连续波形), 传统的 YOLOv11 模型中感受野是固定的, 卷积操作难以同时捕获细粒度频率和长时域波形特征, 信号在频率轴需要高分辨率特征提取, 在时间轴需要长程依赖建模能力, 常规卷积的单一感受野难以兼顾二者, 因此, 本文在单一感受野基础上设计一种能够兼顾二者的三重感受野模块。在传统 YOLOv11 中 C3k2 模块结构是: 一个  $1 \times 1$  卷积降低通道数; 然后通过多个 Bottleneck (每个 Bottleneck 包含两个  $3 \times 3$  卷积) 进行特征提取; 最后再通过  $1 \times 1$  卷积调整通道数。每个  $3 \times 3$  卷积层的感受野大小是固定的 (例如, 一个  $3 \times 3$  卷积在  $stride=1$  且  $padding=1$  的情况下, 感受野为  $3 \times 3$ )。当多个卷积层堆叠时, 整个模块的感受野会随着深度增加而增大, 但是在同一层次上, 每个模块内部的感受野是固定的, 无法同时捕获信号时频图多尺度特征信息, TRF 模块通过使用多个分支 (每个分支有

不同的膨胀率) 实现多尺度感受野, 使的同一层特征图能够同时捕获不同尺度的上下文信息。

TRF 结构如图 4 所示, TRF 模块有三个分支, 分别使用不同的膨胀率 (dilation) 和填充 (padding) 来构建不同感受野。 $D=1$  为小感受野分支, 使用较小的膨胀率和填充, 捕获信号时频图局部细节。 $D=2$  为中等感受野分支, 使用中等膨胀率和填充, 捕获信号时频图中等范围特征。 $D=3$  为大感受野分支, 使用较大的膨胀率和填充, 捕获信号时频图全局特征。信号时频图中信号特征包括信号的频率分布、信号的时间持续性以及信号的时频模式, 这些信号特征以不同的尺度出现, 例如, 信号在频率轴上的覆盖范围广 (大尺度特征), 而一个脉冲信号在时间轴上很短 (小尺度特征)。因此, 使用 TRF 模块可以同时关注这些不同尺度的特征, 增强模型对多尺度特征感知能力, 从而提高模型对信号识别准确率。

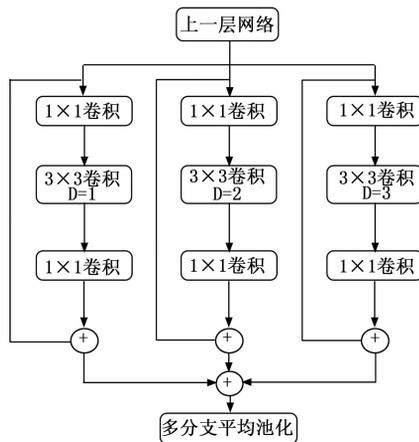


图 4 TRF 结构图

TRF 模块中每个分支的感受野计算公式如下所示:

$$RF_L = RF_{L-1} + (K_{eff} - 1) \cdot S_{total} \quad (3)$$

$$K_{eff} = K + (K - 1) \cdot (D - 1) \quad (4)$$

式中,  $RF_L$  为第  $L$  层的输出感受野,  $RF_{L-1}$  为第  $L-1$  层输入感受野,  $K_{eff}$  为等效卷积核大小,  $S_{total}$  表示之前所有层步长 (stride) 的累计乘积,  $K$  表示原始卷积核大小,  $D$  表示膨胀率。

### 3.3 损失函数改进

损失函数是一个数学函数, 通过模型的预测输出与真实标签作为输入, 计算出表示模型性能的数值, 用来衡量预测模型与真实值之间的差距, 损失函数越小代表模型置信度越高。在目标检测识别任务中, 边界回归损失函数对模型性能具有决定性作用, 现有基于交并比 (IoU, intersection over union) 的损失函数存在不合理的惩罚因子, 导致锚框在回归过程中出现面积扩大的问题, 使得回归效率低, 收敛速度变慢。同时, 这些损失函数在反应锚框与目标之间的差异、考虑目标锚框大小

以及某些情况下的性能表现等方面存在局限性。为解决上述问题,本文设计使用 PIoUv2 损失函数, PIoUv2 损失函数设计了以目标框大小为分母的自适应惩罚因子:

$$P = \left( \frac{dw_1}{w_{gt}} + \frac{dw_2}{w_{gt}} + \frac{dh_1}{h_{gt}} + \frac{dh_2}{h_{gt}} \right) / 4$$

式中,  $dw_1$ 、 $dw_2$ 、 $dh_1$ 、 $dh_2$  是预测框与目标框对应边距离绝对值,  $w_{gt}$ 、 $h_{gt}$  是目标框的宽和高。这样设计使得惩罚因子仅取决于目标框大小,避免了锚框扩大。通过引入基于锚框质量的梯度调整函数  $f(x) = 1 - e^{-x^2}$ , PIoU 损失函数定义为:

$$L_{Piou} = 1 - IoU = L_{IoU} + f(P)$$

其范围是  $0 \leq L_{Piou} \leq 2$ 。损失函数能够根据锚框质量自适应调整梯度大小,引导锚框更直接高效地回归。为避免模型出现过拟合风险 PIoUv2 损失函数在 PIoU 损失函数基础上引入由单个超参数  $\lambda$  控制的非单调注意力层。定义了注意力函数  $q = e^{-p}$ 、 $u(x) = 3x \cdot e^{-x^2}$ , PIoUv2 损失函数为:

$$L_{PiouV2} = 1 - u(\lambda q) \cdot L_{Piou} = 1 - 3 \cdot (\lambda q) \cdot e^{-(\lambda q)^2} \cdot L_{Piou}$$

当  $A(P) = 3\lambda e^{-p} e^{-(\lambda p)^2}$ 、 $B(P) = (1 - IoU + e^{-P^2})$  时,损失函数可表示为  $L_{PiouV2} = 1 - A(P) \cdot B(IoU, P)$ , 其梯度公式如下所示:

$$\frac{\partial L_{PiouV2}}{\partial b} = - \left[ \frac{\partial A}{\partial b} \cdot B + A \cdot \frac{\partial B}{\partial b} \right] \quad (8)$$

式中,  $A(P)$  由指数函数组成,是无限次可微的,保证了梯度的平滑性,  $B(IoU, P)$  也是平滑函数,因此整体损失函数的梯度是平滑的,当  $IoU=1$  时候,  $L_{Piou} = 0$ ,  $L_{PiouV2} = 1 - 3\lambda$  调整参数可使其为 0,证明了损失函数的无偏性。PIoUv2 损失函数增强了模型对中等质量锚框的关注能力,提高模型定位精度,从而提高模型对信号检测识别能力。

## 4 实验设计与结果分析

### 4.1 数据集构建

为了测试改进后 YOLO 模型短波信号识别效果,本文使用根据短波信号协议仿真生成的数据集进行信号识别实验验证。根据短波协议对仿真生成的短波信号数据集进行分类,构建模型训练所需要的数据集格式,具体流程如下。

步骤 1: 根据短波信号协议仿真生成 2GALE、3GALE、CLOVER2000、CIS12、STANAG4285 和 LINK11 CLEW 这 6 类短波信数据。

步骤 2: 对生成的短波信号仿真数据集加入高斯白噪声, SNR 范围为覆盖实际侦察场景的典型值  $-6 \sim -12$  dB。

步骤 3: 对不同信噪比下的 6 类短波信号进行 STFT 变换,生成训练所需信号时频图数据集,每类信号时频图数量为 2 000 张。

步骤 4: 将信号时频图每一类按照 7 : 3 比例随机分为训练集与验证集,训练集短波信号时频图有 8 400 张,测试集短波信号时频图有 3 600 张,六类信号共计 12 000 张。短波信号数据集具体参数如表 1 所示。

表 1 短波信号数据集参数

短波信号类型	调制方式	中心频率 /Hz	信号带宽 /Hz	符号速率 /Bd
2GALE	8FSK	2 600	2 400	125
LINK11 CLEW	DQPSK	2 600	2 500	45.45
CIS12	12 音	2 600	3 100	120
CLOVER2000	8 音	2 600	2 000	62.5
STANAG4285	8PSK	2 600	2 750	2 400
3GALE	8PSK	2 600	3 000	2 400

### 4.2 实验过程与结果分析

本模型对数据集训练使用 NVIDIA GeForce RTX6000 GPU、CUDA11.8 进行训练,选用 Python 编程语言,调用 Pytorch 深度学习库完成模型构建,在训练批量为 [64, 128, 256] 和初始学习率为 [0.01, 0.02, 0.05] 的预定义空间进行 50 次随机采样实验,最终结果表明,批量大小为 128,初始学习率为 0.01 的组合在验证数据集上实现了最佳准确率与训练效果的平衡,因此本实验训练批量大小为 128,初始学习率设置为 0.01,使用 Adam 优化器,训练过程采用余弦退火衰减,为防止模型过拟合,权重衰减系数设置为 0.000 5,训练总周期为 100 次。

YOLOv11 模型与基于 YOLOv11 改进后的 WT-YOLO 模型对 6 类短波信号平均识别准确率结果如图 5 所示,根据实验结果可知,当训练周期到 25 次左右时改进后的 WT-YOLO 模型平均识别准确率达到 96.4%,当训练周期达到 64 次左右时,传统 YOLOv11 模型识别准确率达到 91.9%,改进后的 WT-YOLO 模型相较于传统 YOLOv11 模型架构训练周期更短,且对短波信号数据集识别准确率更高。

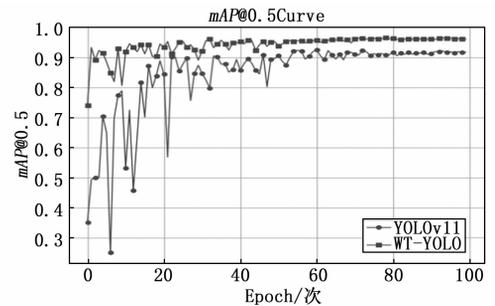


图 5 YOLOv11 与 WT-YOLO 识别结果

查准率 (Precision) 也称为精度,是评估模型性能的一个重要指标,其计算公式如下所示:

$$Precision = \frac{TP}{TP + FP}$$

其中:  $TP$  为真正例, 反映模型正确识别正类样本的能力,  $FP$  为假正例, 反映模型误检情况。查全率 ( $Recall$ ) 也是评估模型性能指标之一, 它衡量的是所有实际为正类样本中被模型正确预测为正类的样本所占的比例, 其计算公式如下所示:

$$Recall = \frac{TP}{TP + FN}$$

其中:  $TP$  为真正例,  $FN$  为假负例, 即模型错误地将正类样本判断为负类。WT-YOLO 模型查准率—查全率曲线图如图 6 所示, 曲线与坐标轴包围的面积表示该类信号平均识别准确率, 分析结果图像可知, CIS12 信号与 STANAG4285 信号识别准确率为 99.5%, 2GALE 信号识别准确率为 95.6%, 3GALE 信号识别准确率为 91.4%, LINK11 CLEW 信号识别准确率为 94.1%, CLOVER2000 信号识别准确率为 98.5%, 以上结果说明改进后的 WT-YOLO 模型能够对 6 种信号进行有效识别。

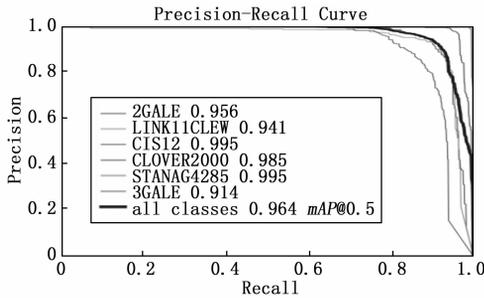


图 6 WT-YOLO 查准率—查全率曲线

为验证 WT-YOLO 模型的优势, 将 WT-YOLO 模型架构与当前常用信号检测识别模型 YOLOv5、YOLOv8 和 YOLOv11 进行实验对比, 所有模型均采用一样的实验配置和相同的数据集, 实验结果如表 2 所示。

表 2 不同网络模型对比

Models	Precision /%	Recall /%	mAP@0.5	Parameters /M	GFLOPs
YOLOv5	90.1	83.6	88.1	2.5	7.1
YOLOv8	92.6	85.3	88.7	3.2	8.7
YOLOv11	93.4	88.4	91.9	2.6	6.3
WT-YOLO	97.2	92.6	96.4	3.0	5.9

通过分析模型对比实验结果可知, WT-YOLO 模型检测识别结果在不同指标上均有所提升, 其中  $mAP@0.5$  相较于 YOLOv5、YOLOv8 和 YOLOv11 分别提升 7.3%、6.7% 和 4.5%。虽然 WT-YOLO 模型参数量较 YOLOv11 增加 0.4 M parameters, 但其运算次数较 YOLOv11 减少 0.4 G 次, WT-YOLO 相较于传统检测识别模型有着更快的模型运算速度, 且相较于对比中的

几个模型有着更高的识别准确率, 有效的解决了短波信号检测识别精度不高、计算复杂度高等问题, 同时有着更低的运算次数, 保证了短波信号检测识别的高实时性。

为进一步验证模型在工程实际应用中队短波信号检测识别性能, 本文采用经典短波 Watterson<sup>[23]</sup> 信道模型, 仿真了低纬度静态、低纬度中等、低纬度恶劣、中纬度静态、中纬度中等、中纬度恶劣、高纬度静态、高纬度中等等 8 种信道环境参数下信号的传输过程, 不同短波信道环境具体参数如表 3 所示。

表 3 不同短波信道环境参数

信号模型	时延扩展/ms	多普勒扩展/Hz
低纬度静态环境	0.5	0.5
低纬度中等环境	2	1.5
低纬度恶劣环境	6	10
中纬度静态环境	0.5	0.1
中纬度中等环境	1	0.5
中纬度恶劣环境	2	1
高纬度静态环境	1	0.5
高纬度中等环境	2	10

不同信道环境下测试结果如图 7 所示, 根据实验结果分析可知, 虽然各种信道环境下模型识别准确率存在差异, 但当  $SNR > 10$  dB 时, 本文所提算法在 8 种信道环境下的平均识别准确率达到 96% 左右, 进一步验证 WT-YOLO 模型架构对信号检测识别的有效性。

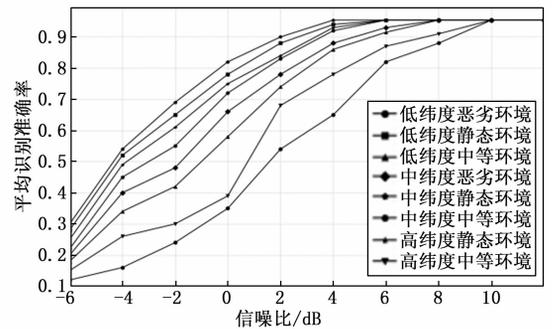


图 7 Watterson 信道下 WT-YOLO 信号识别率

为验证改进的各个模块对模型提升能力的影响, 本文设计了消融实验对各个模块改进效果进行分析, 首先使用小波变换上采样 WFU 模块替换原模型中最邻近插值上采样模块, 提高模型对高频特征与低频特征融合能力, 其次, 使用设计三重感受野 TRF 模块替代 C3k2 模块, 提高模型对信号不同特征的提取能力, 模型消融实验结果如表 4 所示, 分析可知, 引入 WFU 模块后增强了网络模型对信号时频图的高频特征和低频特征提取, 网络模型虽然参数量有所提升, 但其识别准确率较传统 YOLOv11 模型提高了 1.5%, 引入 TRF 模块提高了网络模型捕获时频图多尺度特征能力, 其参数量相较于传

统 YOLOv11 降低了 0.3 M, 计算量降低 1.2 G 次运算, 识别准确率提高 2.9%, 整体 WT-YOLO 模型识别准确率提高 4.5%, 虚警概率相对降低了 55.6%, 其计算公式如下所示:

$$\frac{FP_{YOLOv11} - FP_{WT-YOLO}}{FP_{YOLOv11}} \times 100\% = 55.6\% \quad (11)$$

证明了改进后模型较原始模型有更高的识别准确率。

表 4 消融实验分析

Models	mAP@0.5	Parameters/M	GFLOPs
YOLOv11	91.9	2.6	6.3
YOLOv11+WFU	93.4	3.6	8.1
YOLOv11+TRF	94.8	2.3	5.1
YOLOv11+WFU+TRF+PIoUv2	96.4	3.0	5.9

各个模块对不同信号类别识别结果如图 8 所示, 图 8 (a) 中混淆矩阵表示未改进前模型对 6 种信号识别结果, 图 8 (b) 中混淆矩阵表示引入 WFU 模块后模型对 6 种信号识别结果, 以 CLOVER2000 信号为例, 其信号由 8 个时域脉冲构成, 边缘清晰但容易混叠而模糊, 引入 WFU 模块后, 其识别准确率由 93.8% 提升至 95.4%, 虽然参数量增加 1 M, 但是 WFU 模块更好地保留了信号的纹理细节。图 8 (c) 是在传统模型中引入

TRF 模块后各种信号识别准确率结果图, 以 LINK11CLEW 为例, 引入 TRF 模块后, 其识别准确率由 91.1% 提升至 92.5%, 且参数量减少 0.3 M, TRF 模块的引入提高了模型对多尺度特征敏感的信号识别精度。图 8 (d) 是本文最终模型 WT-YOLO 的信号识别结果图, 对比其它三张混淆矩阵图可知, 最终模型 WT-YOLO 在总参数量仅增加 0.4 M 的情况下, 实现了 4.5% 平均识别准确率的提升, 且每种信号识别准确率较原始模型都有提升, 证明了设计的模块在参数利用上的高效性, 成功地在模型性能与复杂度之间取得平衡。

### 5 结束语

本文针对短波信号识别准确率低的问题, 分析研究了短波信号在时频图上呈现的视觉特征, 提出了一种基于改进 YOLOv11 的深度学习网络模型架构, 通过引入基于小波变换上采样 WFU 模块、设计三重感受野 TRF 模块和损失函数 PIoUv2 模块, 提高模型特征融合能力和捕获信号时频图多尺度特征能力, 从而提高模型检测识别精度。三重感受野 TRF 的设计使得模型训练时间较原始 YOLOv11 降低一倍左右, 在保证高识别准确率情况下, 最大程度减少模型训练所需时间。实验结果表明, 改进后的 WT-YOLO 较 YOLOv11 模型对短波信号具有更高的识别准确率, 证明了该模型在短波信号侦察识别中的有效性。

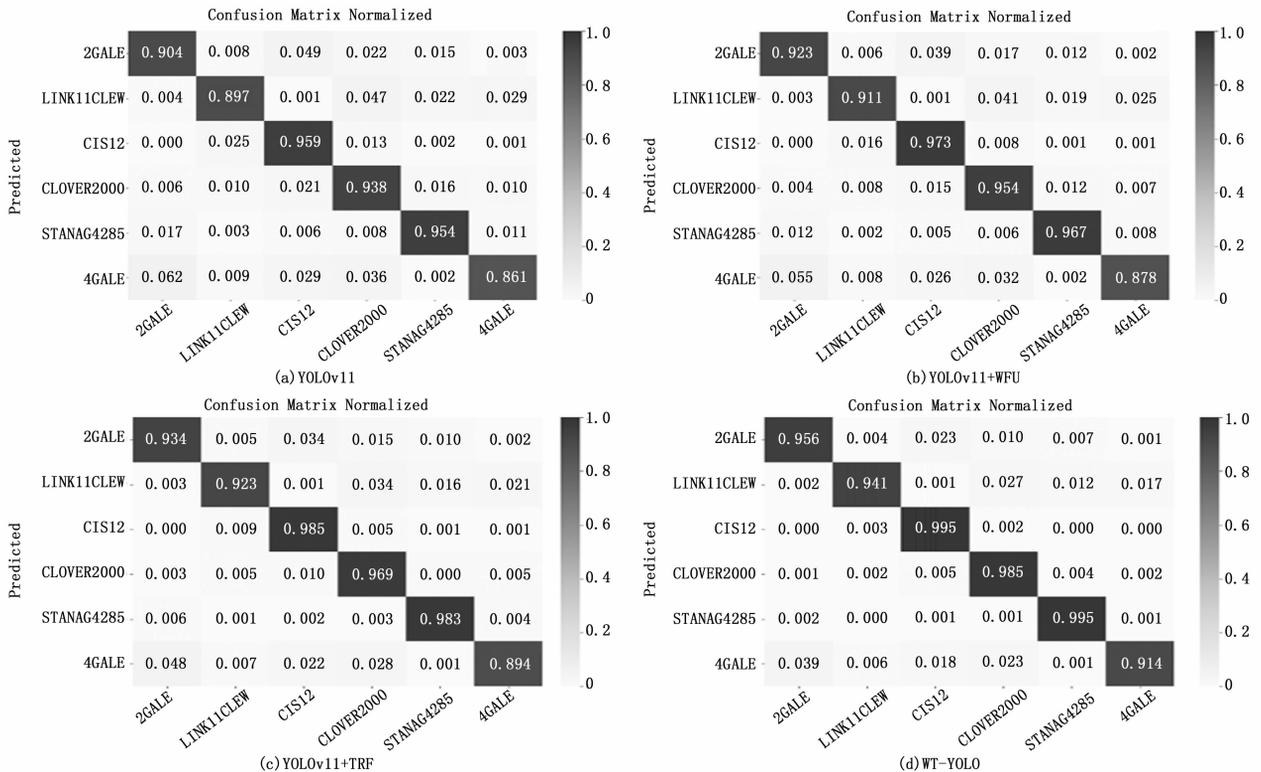


图 8 不同网络模型混淆矩阵

(下转第 165 页)