

改进 YOLOv5 的闪电哨声波轻量化 自动检测模型*

路超^{1,2} 泽仁志玛^{1,2} 杨德贺² 孙晓英²
吕访贤² 冉子霖^{1,2} 申旭辉³

1(应急管理部国家自然灾害防治研究院 北京 100085)

2(中国科学院大学应急管理科学与工程学院 北京 100049)

3(中国科学院国家空间科学中心 北京 100190)

摘要 提出一种改进 YOLOv5(You-Only-Look-Once version 5) 检测模型 YOLOv5-Upgraded. 为了更快定位真实边框, 该模型将损失函数 CIOU (Complete IoU) 替换为 SIOU(Scylla IoU); 同时为了避免网络训练过程中梯度消失、梯度爆炸以及神经元坏死等现象, 将激活函数 SiLU(Sigmoid-weighted Linear Unit) 替换为具有更好梯度流的 Mish; 在主干网络中插入注意力 (Coordinate Attention, CA) 机制, 帮助模型更精准地识别闪电哨声波, 大大降低了漏检率. 基于张衡一号感应磁力仪 (Search Coil Magnetometer, SCM) 数据, 以 2.4 s 时间窗口截取数据, 经带通滤波、短时傅里叶变换得到 1126 张时频图数据集, 再经图像增强操作扩充至 7882 张, 其中 7091 张作为训练集, 791 张作为测试集. 实验结果表明, 基于改进 YOLOv5 的模型平均精度均值为 99.09%, 召回率为 96.20%, 与 YOLOv5s 相比, 分别提升了 2.75% 和 5.07%, 与基于时频图的 YOLOv3 模型相比, 平均精度均值和召回率则分别提高了 5.89% 和 9.62%. 基于智能语音的 LSTM (Long Short Term Memory Networks) 闪电哨声波识别模型大小为 82.89 MB, YOLOv5-Upgraded 仅为 13.78 MB, 约节省 83.38% 的内存资源. 研究表明改进后的轻量化模型大大降低了闪电哨声波的漏检现象, 在测试集中取得了较好结果, 并且其轻量化特征易于部署到卫星设备, 极大地提高了星载识别的可能性.

关键词 张衡一号卫星, 闪电哨声波, YOLOv5, 轻量化, 自动检测

中图分类号 P352

Lightweight Automatic Detection Model for Lightning Whistle Waves Based on Improved YOLOv5

LU Chao^{1,2} ZEREN Zhima^{1,2} YANG Dehe² SUN Xiaoying²
LÜ Fangxian² RAN Zilin^{1,2} SHEN Xuhui³

1(National Institute of Natural Hazards, Ministry of Emergency Management of China, Beijing 100085)

* 国家自然科学基金面上项目资助 (41874174)

2023-06-12 收到原稿, 2023-08-26 收到修定稿

©The Author(s) 2024. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

2(School of Emergency Management Science and Engineering, University of Chinese Academy of Sciences, Beijing 100049)

3(National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

Abstract This project proposes an improved YOLOv5 detection algorithm YOLOv5 Upgraded. To address this issue, the study proposes an improved YOLOv5 detection algorithm called YOLOv5-Upgraded. The model takes into account the vector angle between the predicted edge and the real edge, The model replaces the loss function CIOU (Complete IoU) with SIOU (Scylla IoU); at the same time, in order to avoid phenomena such as gradient disappearance, gradient explosion, and neuron necrosis during network training, the activation function SiLU (Sigmoid-weighted Linear Unit) is replaced with Mish with better gradient flow; The CA attention mechanism is inserted into the backbone network to help the model identify the Lightning whistler waves more accurately and greatly reduce the missed detection rate. The study is based on the VLF-band data of CSES Satellite SCM with 2.4 seconds time window to intercept data, and 1126 time-frequency map data sets are obtained by band-pass filtering and short-time Fourier transform, and then expanded to 7882 images by image enhancement operations, of which 7091 are used as training set and 791 are used as test set. Experimentally, the average mean accuracy (mAP) of the improved YOLOv5-based model is 99.09% and the Recall is 96.20%, which are improved by 2.75% and 5.07% compared with the plain YOLOv5s, and 5.89% and 9.62% compared with the time-frequency map-based YOLOv3 model. The size of LSTM based on the speech processing technology lightning whistler waves recognition model is 82.89MB, while the YOLOv5-Upgraded model is only 13.78 MB, saving about 83.38% of memory resources. It is shown that the model greatly reduces the leakage problem of Lightning whistler waves, achieves better results in test set, and its lightweight features are easy to deploy to satellite devices, which greatly improves the possibility of satellite recognition.

Key words CSES, Lightning whistler waves, YOLOv5, Lightweight, Automatic detection

0 引言

闪电是地球大气层内频发的自然现象, 平均每秒发生 44 次左右, 会产生覆盖宽频段的电磁波, 由于其高频部分的群速度大于低频部分群速度, 因此在电磁卫星轨道空间内, 闪电引起的宽频带电磁辐射往往高频部分先到达, 随后低频部分才到达, 呈现典型的哨声波特征, 故而称为闪电哨声波^[1]. 当闪电哨声波的传播路径较长、沿着传播路径的电子密度或磁场强度较大时, 呈现较为清晰的下降调, 即 L 形态色散特征^[2].

当前卫星观测的闪电哨声波的数据主要来源于 AKEBONO 卫星、Arase 卫星、DEMETER 卫星以及张衡一号卫星^[3-5]. 张衡一号电磁卫星 (China Seismo-Electromagnetic Satellite, CSES) 于 2018 年 2 月 2 日在酒泉卫星发射中心发射成功, 标志着中国具备了天基观测空间电磁场及等离子体环境的能力. CSES 观测范围覆盖南北纬 65°, 在中国及周边区域进行详查模式的监测, 其余区域为巡查模式. 卫星轨

道倾角为 97°, 重访周期为 5 天, 在轨高度约 507 km, 位于电离层和等离子体层边界.

张衡一号卫星搭载的感应式磁力仪和电场仪实现了变化磁场从 10 Hz 至 20 kHz、空间电场从直流 DC 至 3.5 MHz 的观测, 卫星在轨五年获取了海量的电磁场观测数据^[6,7], 记录了众多类型的电磁波动现象, 例如准周期辐射、磁力线谐振, 以及最为常见的电离层嘶声、闪电哨声波等^[8-10]. 目前对张衡一号卫星的海量电磁场数据的应用研究很少, 传统依靠人工分析的数据处理方式效率较低. 随着大数据分析技术、人工智能等高新信息处理技术的发展, 为高效处理海量电磁场数据带来了契机. 本文拟通过设计智能检测算法来检测闪电哨声波, 以期在较大程度上提高闪电哨声波的检测效率. 类似研究工作已经开展, 例如 Fiser 等^[11]提出了基于交叉熵的闪电哨声波识别算法, 制作了白天和黑夜的闪电哨声波模板, 再利用模板匹配策略识别闪电哨声波, 但是该算法的检测结果受特征影响较大, 对时频噪声十分敏感, 检测效果

不佳; Oike 等^[12]在该基础上制作了不同色散形态的闪电哨声波模板, 提高了识别精度; Ali Ahmad 等^[13]认为以上方式很难检测其他色散特征的闪电哨声波, 用拉普拉斯算子做边缘提取, 提出了能检测出多种色散形态特征提取方法, 再基于决策树分类算法进行识别.

近年来, 深度学习理论快速发展, Konan 等^[14]提出了基于滑动深度卷积神经网络 (Sliding Deep Neural Convolutional neural Network, SDNN) 以及基于 YOLOv3 神经网络的两种闪电哨声波检测方法. SDNN 依靠多个卷积层提取图像特征, 使得模型具有较强的泛化能力, 但基于固定时间长度的策略易导致其他时间长度的闪电哨声波被忽略. 基于 YOLOv3 的闪电哨声波检测算法包含两个部分: 主体网络以及非极大抑制算法 (Non-Maximum Suppression, NMS). 其中主体网络采用了残差结构使得构建深层网络模型的同时避免了梯度消失以及梯度下降问题, 与此同时输出多个粗定位的预测边框, 最终使用 NMS 算法对输出的预测边框进行过滤, 保留最佳边框, 使闪电哨声波的检测效果得到明显改善; 但 YOLOv3 模型所消耗内存较大, 不易于部署星载. Yuan 等^[15,16]提出了两种闪电哨声波的识别方法: 基于支持向量机 (SVM) 的识别算法, 以及基于智能语音技术的长短期记忆网络 (LSTM) 的识别算法. 前者通过设计能增强闪电哨声波 L 形态特征的卷积核以及采用支持向量机作分类器提高了识别精度, 但同样因基于时间窗的方式导致超出该窗口的闪电哨声波被忽略; 后者 LSTM 首次采用智能语音技术进行闪电哨声波的自动识别, 然而其设计基于人耳听觉特性, 在低频段使用了较多数量的滤波器, 在中高频区域使用较少数量的滤波器, 这种设计在提取较低频段的闪电哨声波时效果良好, 但在提取较高频段数据的闪电哨声波时提取精度会降低.

综上所述, 已有的闪电哨声波检测模型存在两个问题, 即无法兼顾检测精度以及模型的内存消耗. 当前基于时频图的 YOLOv3 模型虽然有较高的检测精度, 但是模型内存消耗较大; 基于智能语音技术的 LSTM 模型在 YOLOv3 模型的基础上降低了内存消耗, 但是对于较高频段的闪电哨声波检测精度大大降低, 并且以上两种模型在本文数据集上均存在较为严重的漏检情况, 效果较差, 不在这里讨论. 针对以上问题, 本文基于 YOLOv5 设计检测模型, 在 YOLOv5

模型中增加了大量的残差网络, 具有检测精度更高、模型内存消耗更小的特点, 适用于闪电哨声波的检测.

1 数据预处理

1.1 原始数据集制作

张衡一号卫星搭载的感应式磁力仪能够提供从超低频到甚低频三个频段的频谱和波形数据. 其中超低频 ULF (Ultra Low Frequency) 波段 10~200 Hz, 极低频 ELF (Extremely Low Frequency) 波段 0.2~2.2 kHz, 甚低频 VLF (Very Low Frequency) 波段 12.5~20000 Hz^[17]. 本研究主要采用了 SCM 在 VLF 频段的波形数据, 数据采集时间范围为 2019 年 11 月至 2020 年 4 月. 图 1 显示了张衡一号卫星在 VLF 频段记录的闪电哨声波事件.

从图 1 可以看出, 在 1~6 kHz 上出现了典型的闪电哨声波, 其时频呈现出明显的下降调特征, 表现为 L 形态. 张衡一号卫星常年记录的多个闪电哨声波的频谱特征及频率范围如图 1 所示, 本研究则是基于 1~6 kHz 频段的波形数据进行的. 为检测到持续时间更长的闪电哨声波开展试验, 以 2.4 s 时间窗口截取波形数据, 对截取数据进行截止频率为 1~6 kHz 的带通滤波处理, 通过短时傅里叶变换 (FFT) 转换成时频图, 结果如图 2 所示. 图 2 为经带通滤波处理前后波形数据与时频特征的比较, 图 2(a) 为原始波形, 图 2(b) 为 FFT 转换后的时频结果, 图 2(c) 为原始波形数据经带通滤波过滤后的波形数据, 其 FFT

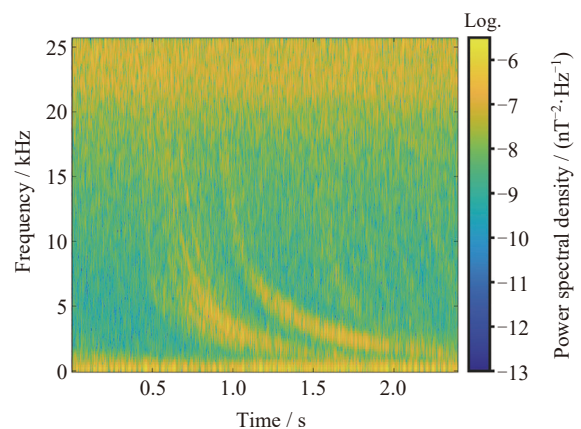


图 1 张衡一号卫星感应式磁力计记录的闪电哨声波事件 (PSD 为功率谱密度)

Fig. 1 Whistler wave event recorded by SCM onboard CSES satellite (Power Spectral Density, PSD)

转换后的时频结果如图2(d)所示. 比较发现, 经过带通滤波器之后得到的闪电哨声波频谱特征更加明显.

由于训练过程中模型的输入特征图仅需要展示原始时频图信息以及标记的闪电哨声波的标签(label), 因此在数据集中无需绘制横纵坐标、色阶栏等坐标信息, 原始特征如图3所示. 这里采用了LabelImg图像数据标注工具在时频图中对L形色散特征绘制矩形外框, 进行闪电哨声波的人工标注, 并生成相应txt格式标签, 供模型训练时读取. 标签中包含五个数据, 分别代表标签类别、归一化之后的中心点横坐标、纵坐标、目标宽度以及目标高度.

1.2 数据增强

模型训练时, 如果数据集的数量较小会产生较大误差. YOLOv5s模型采用了与YOLOv4模型相同的马赛克(Mosaic)图像增强方法, 其主要思想是通过缩放、随机裁剪、拼接操作将四张特征图整合为一张特征图作为训练数据, 丰富了图片背景, 一定程度上增强了模型的泛化能力. 对原始数据进行数据增强, 可以平滑输入空间结构, 提高模型从输入空间学习映射规则的能力, 同时起到扩充数据集的效果. 同时, 由于原始波形数据采集环境是极其复杂的, 例如太阳辐射抛射出高能粒子对载荷造成的硬损伤, 导致原始数

据存在一定误差. 为了提高模型的泛化能力和容错率, 研究基于python-skimage图像处理库对原始数据集进行图像增强工作, 包括亮度增强、亮度减弱、增加椒盐噪声, 此外, 考虑到张衡一号卫星本身在SCM载荷采集数据过程时所产生的噪声, 为图像增加方差为0.02的高斯噪声, 增强后的数据集示例如图4所示, 以下为详细过程.

第一步 图像亮化和图像暗化: 伽马校正(Gamma correction)是用来对图像数据做非线性运算的算子, 可以调节图像的灰度值明暗范围. 当 $\gamma > 1$ 时, 对输入特征图起到暗化作用; $\gamma < 1$ 时, 对输入特征图起到亮化作用. 研究对原始数据集进行 γ 校正, 采用的 γ 分别为1.5, 2.0, 0.8, 0.5得到灰度值不同的4664张输入特征图, 如图4(a)~(d)所示.

第二步 增加椒盐噪声: 椒盐噪声(Salt and pepper noise)是数字图像中经常出现的噪声, 其成因是图像信号突然受到脉冲干扰, 在图像上产生随机的黑点或者白点. 研究对原始数据集增加椒盐噪声, 得到1126张特征图, 使构建的数据集更具有普适性, 有助于提高模型的泛化能力, 如图4(e)所示.

第三步 增加高斯噪声: 高斯噪声或者白噪声(Gaussian Noise)是指概率密度函数符合高斯正态分

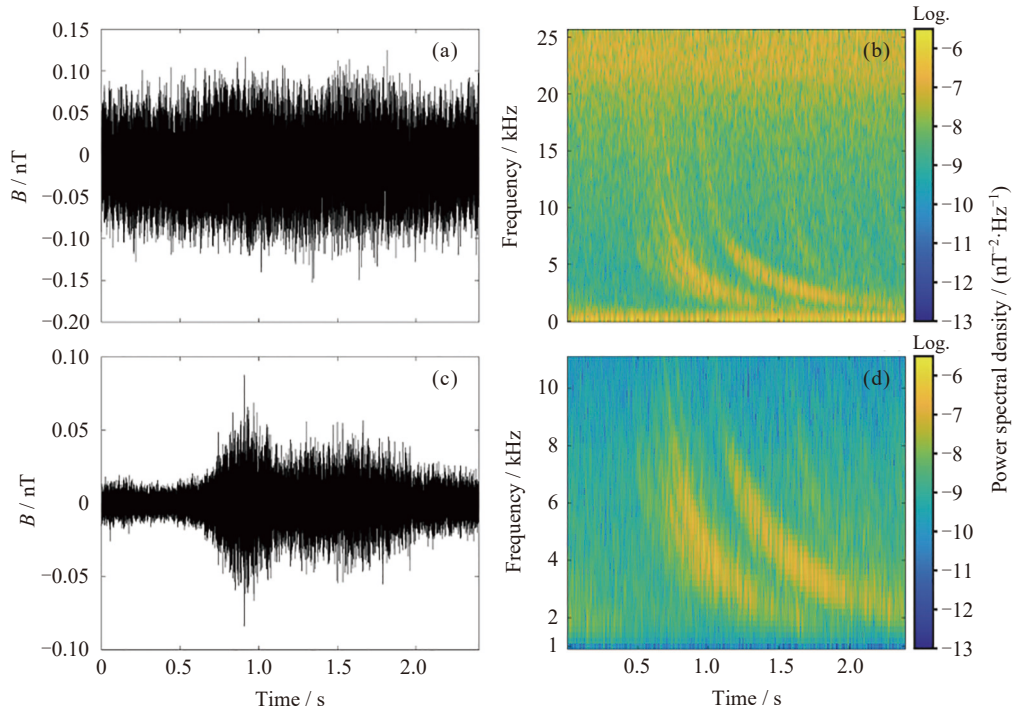


图2 VLF波形以及对应的时频结果

Fig. 2 VLF wave and corresponding time-frequency diagram

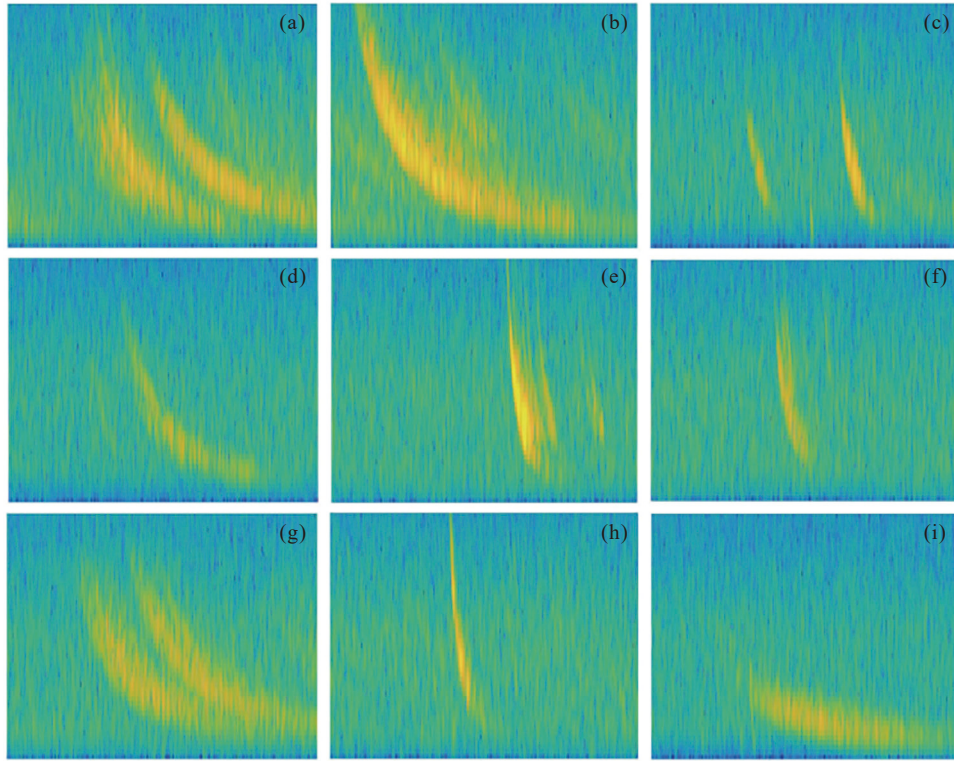


图 3 训练模型输入的原始特征示例

Fig. 3 Examples of origin dataset for the training model

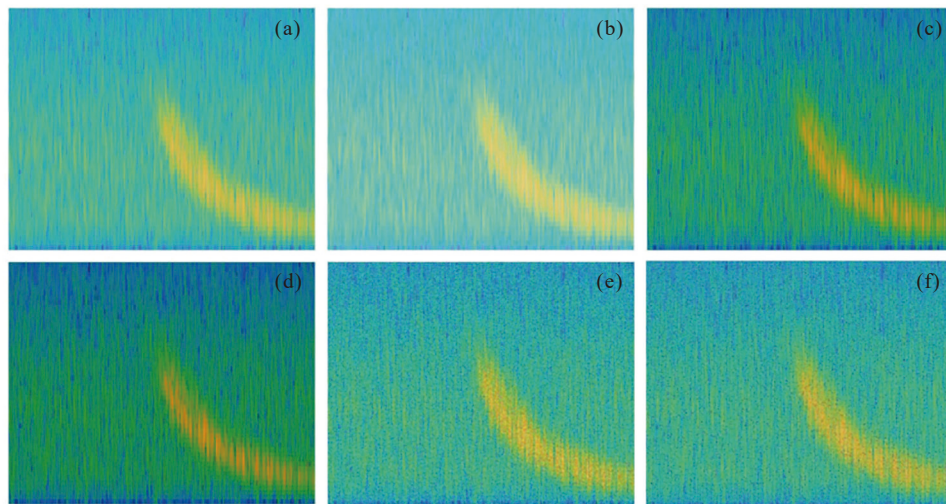


图 4 增强后的数据集示例

Fig. 4 Example of enhanced dataset

布的噪声, 概率密度函数如下:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right]. \quad (1)$$

考虑到张衡一号卫星本体运行时对空间电场和磁场产生的变化^[18], 会对 SCM 载荷采集的数据造成

一定干扰, 本研究对原始数据集增加方差为 0.02 的高斯噪声, 得到 1126 张特征图, 如图 4(f) 所示。

数据集制作流程如图 5 所示。首先以步长为 0.24 s、时间窗为 2.4 s 对原始波形数据进行截取, 经带通滤波过滤保留得到 1~6 kHz 的波形数据, 以

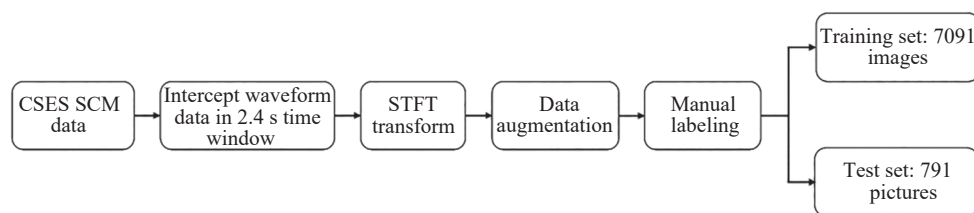


图5 数据集制作流程

Fig. 5 Process of original dataset production

256 的采样点对截取后的波形数据进行短时傅里叶变换得到时频图. 人工筛选后得到 1126 张原始数据集, 经图像增强操作扩充至 7782 张, 并对数据集进行人工标记, 划分为训练集和测试集.

数据集的划分详情如表 1 所示, 其中 7091 张作为训练集, 791 张作为测试集. 训练集用于神经网络模型的训练以及特征提取, 测试集用于对模型测试, 评估模型泛化能力.

2 YOLOv5-Upgraded 轻量化自动检测模型构建

2.1 YOLOv5 算法

YOLO 算法 (You Only Look Once, YOLO) 于 2016 年提出, 相比其他目标检测算法, 该算法在检测精度方面优势明显, 且具有更快的检测速度^[19]. YOLO 系列提出以来经不断更新, 已趋于成熟, 被广泛应用于目标检测领域. 2020 年 Ultralytics 公司发布基于 PyTorch 深度学习框架的轻量级神经网络 YOLOv5, 与之前版本相比, 该版本具有模型权重更小、训练时间更短、推理速度更快的特点. 根据模型参数以及大小不同, YOLOv5 有五个版本, 分别是 YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, 五种版本的模型宽度以及深度依次增大、参数量依次增加, 检测精度上升的同时也增大了模型权重, 本研究旨在保证精度的同时实现闪电哨声波检测模型的轻量化, 并且降低闪电哨声波的漏检率, 针对 YOLOv5s 模型进行改进.

YOLOv5 网络结构包括四个部分, 分别是 Input, Backbone, Neck, Head 四个部分, 网络结构如图 6 所示.

Input 端采用与 YOLOv4 一样的马赛克 (Mosaic) 数据增强方式, 完成图片的自适应随机缩放、锚框计算等基本任务, 提高了对小目标的检测精度.

表 1 数据集划分详情
Table 1 Dataset division details

数据集	训练集	测试集	合计
原始数据	1013	113	1126
图像亮化	2026	226	2252
图像暗化	2026	226	2252
增加高斯噪声	1013	113	1126
增加椒盐噪声	1013	113	1126
合计	7091	791	7882

Backbone 是 YOLOv5 的主干网络, 包含 Focus 操作、卷积层、C3 以及 SPPF 结构. 其中 Focus 操作对原始图像进行隔列采样降低特征图分辨率, 并通过一个包含 32 个卷积核的卷积操作增加输入特征图通道数. 卷积层中包含三个功能, 普通卷积 Conv、数据归一化 (Batch Normalization, BN) 以及激活函数 (SiLU); C3 为 Backbone 的另一关键结构, C3 在 BottleneckCSP 基础上减少一个卷积层, BottleneckCSP 和 CSP 结构相似, CSP 模块将基础层的特征映射划分为两部分, 通过跨阶段层次结构进行合并, 在减少参数量的同时确保检测精度. SPPF 结构为改进的空间金字塔池化, 解决了网络重复提取特征的问题, 加快前向传播和后向传播过程.

Neck 部分采用路径聚合网络 (Path Aggregation Network, PAN) 和特征金字塔 (Feature Pyramid Networks, FPN) 结构, 融合 Backbone 部分提取到的特征, 使模型提取到更加丰富的特征, 加强了网络对不同尺寸对象特征融合的能力.

Head 部分使用损失函数计算模型对闪电哨声波的预测边框与真实边框的交并比 (Intersection over Union, IoU) 作为损失值, 选出 IoU 值最小的检测框作为结果. 在闪电哨声波检测的后处理过程中, 采用非极大值抑制 (Non Maximum Suppression, NMS)

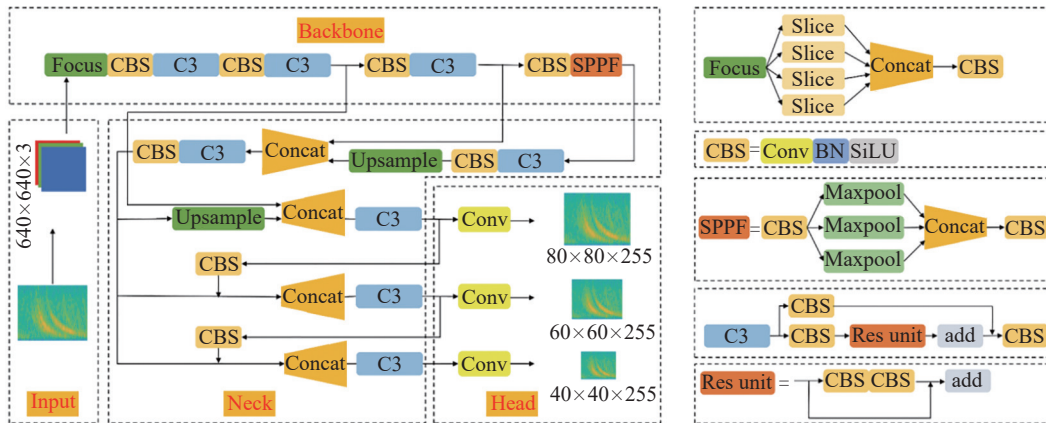


图6 YOLOv5 网络结构示例

Fig. 6 Example of the network structure of YOLOv5

方法增强对密集目标的筛选能力.

2.2 激活函数改进

模型中引入激活函数的目的是添加非线性关系, 保证多层神经网络在训练过程中不退化成单层神经网络. 较多使用的激活函数包括 Sigmoid, ReLU, SiLU, Mish 等.

YOLOv5 中默认使用 SiLU 作为激活函数, 其无上界、有下界以及非单调的特性, 使得其效果比 Sigmoid, ReLU 等激活函数要好, 但存在梯度消失问题. Mish 函数于 2019 年提出, 其是一条非单调的光滑激活函数, 这种特征使其具有较好的稳定性, 允许信息更好地深入模型, 在深层次网络中具有较强的泛化能力^[20]. 在负半轴上 Mish 函数的导数不为 0, 有效解决了训练过程中神经元死亡的现象, 即对权重更新为负数的神经元仍然可以进行更新. 由于 Mish 函数的非饱和特性, 不具有上界, 有效避免了梯度消失问题. 并且 Mish 函数的导数在正半轴上绝大时刻小于 1, 不会出现梯度爆炸问题. Mish 激活函数的图像以及相应导数的图像如图 7 所示, 其公式如下:

$$\begin{aligned} F(x) &= x \tanh[f(x)], \\ \tanh(x) &= \frac{e^x - e^{-x}}{e^x + e^{-x}}, \\ f(x) &= \ln(1 + e^x). \end{aligned} \quad (2)$$

将 YOLOv5 结构中 CBS 结构的 SiLU 激活函数更改为 Mish 后, CBL 结构变为 CBM, 如图 8 所示.

2.3 损失函数改进

模型训练过程中, 损失函数的作用是监督模型的学习进程, 通过对比网络模型每次前向计算结果与真实值的差距, 计算出模型的损失, 进而指导模型下一

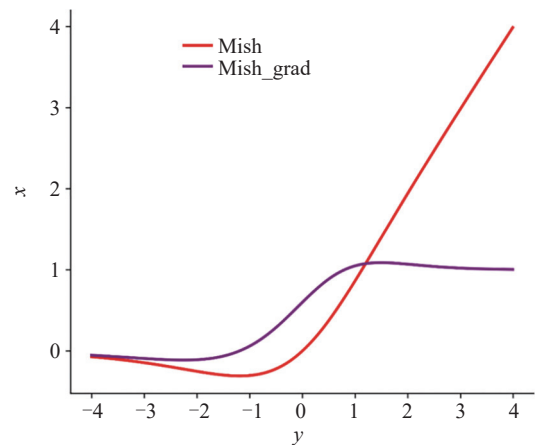


图7 Mish 函数图及其导函数

Fig. 7 Mish function and its derivative function

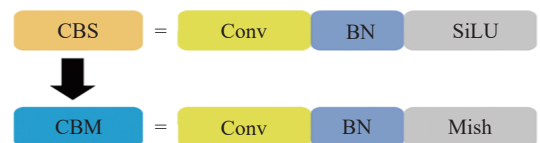


图8 模型结构中 CBS 模块改为 CBM 模块

Fig. 8 Change the CBS module to CBM module in the model structure

步的训练朝着正确的方向进行. YOLOv5 模型 Head 检测头部分共有三种损失函数, 包括分类损失函数 (Classification loss)、目标损失函数 (Object loss) 和边框损失函数 (Bounding box loss). 预测边框和真实边框的交集与并集的比值称为交并比 (Intersection over Union, IoU, 定义符号 R_{IoU}), 如图 9 所示, 相应公式如下:

$$R_{\text{IoU}} = \frac{B \cap B^{\text{gt}}}{B \cup B^{\text{gt}}}. \quad (3)$$

其中, B 为预测边框, B^{gt} 为真实边框.

最初 YOLO 系列模型采用 IoU 作为边框损失函数, 但 IoU 有以下缺点: (1) 当预测边框与真实边框不重叠时, 无法反映出边框之间的远近, 此时损失函数为 0, 不能继续优化模型; (2) IoU 无法精确反映回归效果. 如图 10 所示, 三种情况 IoU 相等, 但可以看出其检测效果不同, 图 10(a) 的效果最好, 图 10(c) 的最差. 随着目标检测领域的发展, GIoU (Generalized IoU), DIoU (Distance IoU), CIoU, EIoU (Efficient IoU) 和 SIoU 等边框损失函数陆续被提出. CIoU 是 YOLOv5 模型默认的损失函数, 不仅考虑到预测边框和真实边框的重叠面积、中心点距离, 还考虑到长宽比, 使得预测边框更加符合真实边框.

EIoU 在 CIoU 的基础上将长宽比损失项分开, 并引入了 Focal loss 使得模型专注 IoU 较大的预测边框. 但以上损失函数没有考虑到真实边框和预测边框之间不匹配的方向, 这种不足会导致网络模型收敛速度较慢且效率较低, 因为预测边框在训练过程中会

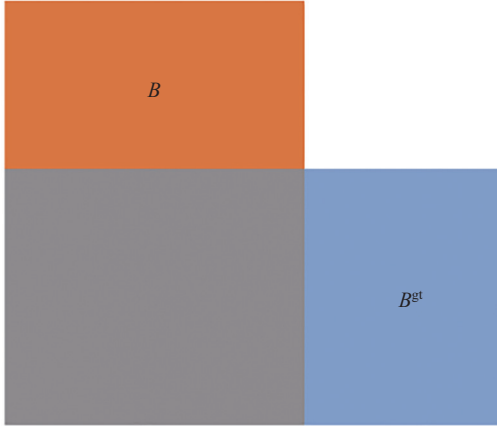


图 9 预测边框和真实边框的 IoU

Fig. 9 IoU of real bounding and predicted bounding

四处游荡导致模型效果更差. Gevorgyan^[21] 提出了一个新的损失函数 SIoU, 考虑传统损失函数惩罚度量之外额外考虑了回归过程中的预测边框与真实边框的向量夹角. SIoU 损失函数由 4 个 Cost 函数组成, 分别是 IoU cost, Angle cost, Distance cost, Shape cost.

研究引入 SIoU 损失函数作为模型 Bounding Box 的定位损失函数, SIoU 使用 Angle cost 的目的是最大限度地减少预测框与真实框距离的变量数量. 如图 11 所示, 模型尝试将预测边框定位到 x 轴或 y 轴, 且沿着相关轴继续定位预测边框.

Angle cost 定义公式如下:

$$\Lambda = 1 - 2 \times \sin^2 \left[\arcsin(x) - \frac{\pi}{4} \right], \quad (4)$$

其中相关变量定义如下:

$$\begin{aligned} \beta &= \frac{\pi}{2} - \alpha, \\ x &= \frac{c_h}{\sigma} = \sin(\alpha), \\ \sigma &= \sqrt{(b_{c_x}^{\text{gt}} - b_{c_x})^2 + (b_{c_y}^{\text{gt}} - b_{c_y})^2}, \\ c_h &= \max(b_{c_y}^{\text{gt}}, b_{c_y}) - \min(b_{c_y}^{\text{gt}}, b_{c_y}). \end{aligned} \quad (5)$$

根据式 (5), Distance cost 被定义为

$$\begin{aligned} \Delta &= \sum_{t=x,y} (1 - e^{-\gamma \rho_t}), \\ \rho_x &= \left(\frac{b_{c_x}^{\text{gt}} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{\text{gt}} - b_{c_y}}{c_h} \right)^2, \gamma = 2 - \Lambda, \end{aligned} \quad (6)$$

可以看出, 当 $\alpha \rightarrow 0$ 时, Distance cost 的贡献大大降低, Shape cost 被定义为

$$\begin{aligned} \Omega &= \sum_{t=w,h} (1 - e^{-\omega_t \theta}); \\ \omega_w &= \frac{|B \cap B^{\text{gt}}|}{\max(w, w^{\text{gt}})}, \quad \omega_h = \frac{|h \cap h^{\text{gt}}|}{\max(h, h^{\text{gt}})}. \end{aligned} \quad (7)$$

其中, θ 的值定义了 Shape cost 及其值是唯一的, 控

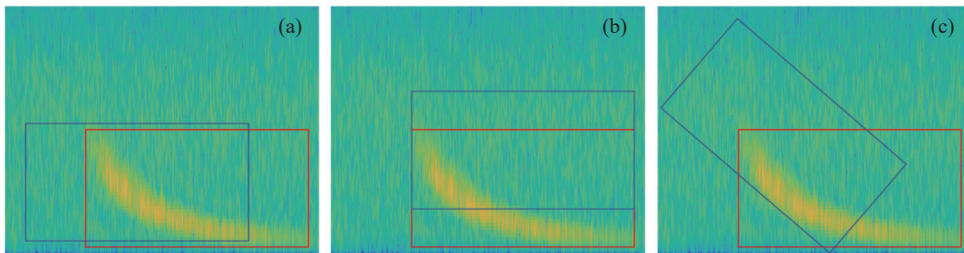


图 10 相同 IoU 对比

Fig. 10 Comparison of the same IoU

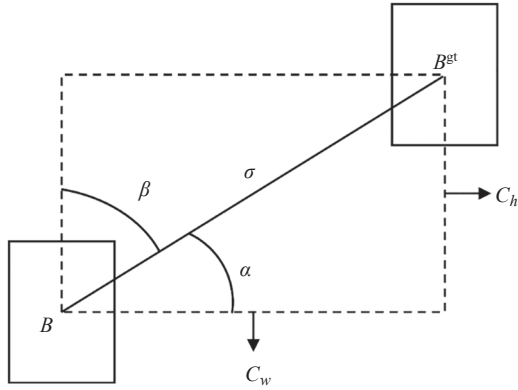


图 11 将角度成本贡献计算到损失函数中的方案

Fig. 11 Scheme for calculation of angle cost contribution into the loss function

制着对 Shape cost 的关注程度. 最终 SIOU loss 如下:

$$L_{\text{box}} = 1 - R_{\text{IoU}} + \frac{\Delta + \Omega}{2}. \quad (8)$$

2.4 引入注意力机制

模型在卷积采样时容易丢失特征图中不够明显的闪电哨声波 L 形态特征信息, 导致检测效果大大降低. 神经网络中引入注意力机制的目的是使模型关注特征图中的更加重要的特征, 从而使模型具有更强的泛化能力.

SE (Squeeze-and-Excitation Network) 注意力机制于 2016 年提出, 该模块用于对模型卷积操作后的特征进行重新校准, 起到增强重要特征、抑制不重要特征的作用^[22]. 但 SE 仅考虑了特征图通道之间的关系, 忽略了位置信息, 之后 CBAM (Convolutional Block Attention Module)、ECA (Efficient Channel Attention) 等注意力机制相继提出, 虽然提高了模型性能但仍仅关注特征图通道之间的关系. CA (Coordinate Attention) 注意力机制于 2021 年提出, 其是一种高效的坐标注意力机制, 不仅关注特征图通道之间的关系, 还考虑到横向和纵向的位置信息, 使模型可以在更大范围上捕获特征^[23].

CA 是一个简单的即插即用的注意力机制, 在提高模型精度的同时几乎不带来额外开销. 本研究在网络中嵌入 CA, 提高模型对闪电哨声波 L 形态特征的提取能力, 图 12 为 CA 原理.

为了获取输入特征图的全局感受野, CA 首先在宽度和高度两个方向上对输入特征图的每个通道进行全局平均池化, 对于输入特征图 X, 其维度为

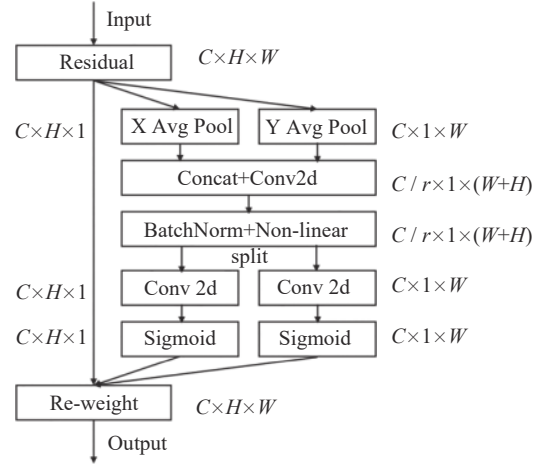


图 12 坐标注意力模块

Fig. 12 Coordinate Attention (CA) model

$C \times H \times W$, 使用尺寸为 $(H, 1)$ 和 $(1, W)$ 的卷积核沿水平方向和垂直方向对通道进行编码, 得到高度为 h 的第 c 个通道和宽度为 w 的第 c 个通道的输出, 有

$$Z_c^h(h) = \frac{1}{w} \sum_{0 \leq i \leq w} x_c(h, i);$$

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq i \leq H} x_c(j, w). \quad (9)$$

以上两个变换沿着两个方向进行特征聚合, 得到两个方向上的特征图 Z^h 和 Z^w , 将其送入一个 1×1 的共享卷积核进行 F_1 变换, 得到一个中间特征, 即

$$f = \delta(F_1([Z^h, Z^w])). \quad (10)$$

沿着空间维度将 f 切分为两个单独的张量, 再利用两个卷积核大小为 1 的卷积运算 F_h 和 F_w 将特征图 f_h 和 f_w 变换为与 X 同样的通道数, 有

$$g^h = \sigma[F_h(f^h)],$$

$$g^w = \sigma[F_w(f^w)]. \quad (11)$$

将 g^h 和 g^w 进行拓展, 作为注意力权重, 最终输出如下:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j). \quad (12)$$

改进后的模型 YOLOv5-Upgraded 的网络结构如图 13 所示, 其中红色箭头标出的为改进后的模块, CBM 模块见图 9.

3 实验结果与分析

3.1 实验环境

研究采用的操作系统型号为 Ubuntu 18.04.6 LTS,

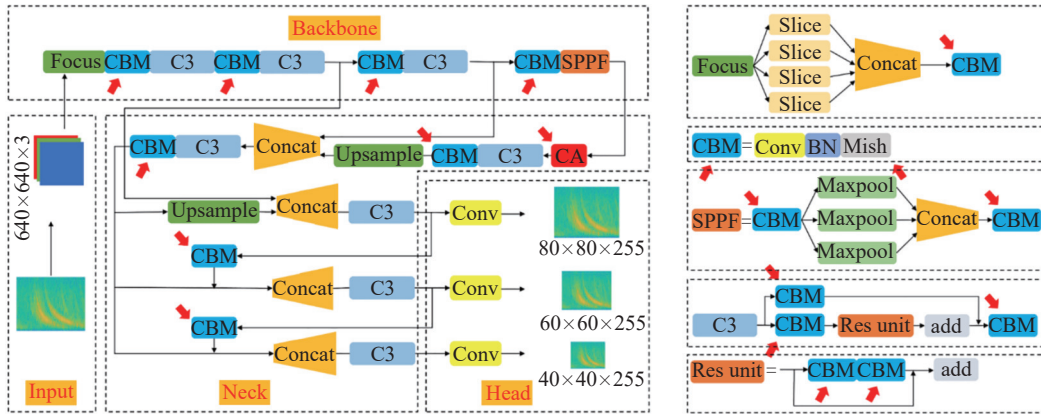


图 13 基于YOLOv5的轻量化自动检测模型(红色箭头标出改进后模块)

Fig. 13 YOLOv5-Upgraded lightweight automatic detection model diagram based on YOLOv5 (The improved module is shown by a red arrow)

CPU 型号为 Intel(R) Xeon(R) CPU E5-2640 v4 @ 2.40 GHz, 运行内存为 256 GB, GPU 型号为 NVIDIA Corporation GV100 GL [Tesla V100 PCIe 32 GB] (rev a1), 内存大小为 32 GB, 为了使模型达到更好性能, 模型训练时设置的超参数列于表 2.

按照以上参数以随机梯度下降 (SGD) 方式进行模型训练, 每次由 BN 层完成正则化以更新模型的权重, 保存训练过程中的最佳权重文件作为最终模型, 在测试集上进行性能评估, 输出闪电哨声波的预测框及置信度.

3.2 评价指标

通过检测精度以及检测速度来评估模型性能. 每秒检测帧数 (FPS) 用作检测速度的指标, 召回率 (Recall)、精确度 (Precision)、F1 值 (F1-Score)、平均精度均值 (mAP) 用来评估模型的精度. 在数据集中 N 表示负类样本, P 表示正类样本, N 与 P 的四种组合及含义如表 3 所示, 当阈值大于 IoU 时, 表示模型预测正确, 当阈值小于 IoU 时, 表示模型预测错误.

定义 TP 为 x , FN 为 y , FP 为 z , TN 为 w , 各公式表示如下.

精确度 (Precision): 精度是模型精确性的度量, 表示被预测为正例 (TP+FP) 的样本中实际正例 (TP) 所占的比例, 该指标越大证明模型精度越高, 有

$$f_p = \frac{x}{z + x}. \quad (13)$$

召回率 (Recall): 召回率是覆盖面的度量, 表示所有正例 (TP+FN) 中实际正例 (TP) 所占的比例, 该指标越大证明模型漏检率越低, 有

表 2 模型训练的超参数设置

Table 2 Hyperparameter setting of model training

参数名	参数值
Weight decay (权重衰减)	0.0001
Momentum (动量)	0.900
Batch size (批量大小)	16
Training epoch (训练轮次)	150
Learning rate (学习率)	0.001

表 3 符号定义

Table 3 Symbol definition

预测类别	描述
TP	正确预测, 被模型预测为正类的正样本
FN	错误预测, 被模型预测为负类的正样本
FP	错误预测, 被模型预测为正类的负样本
TN	正确预测, 被模型预测为负类的负样本

$$f_r = \frac{x}{x + y}. \quad (14)$$

平均精度均值 (mean Average Precision, mAP): mAP 是目标检测领域常见的指标, 越大证明模型平均检测精度越高, 有

$$f_m = \frac{1}{N} \sum_{i=1}^N P(k) \Delta R(k). \quad (15)$$

式中, N 为测试集总样本数量, P 为精确度, k 为第 k 个样本, $\Delta R(k)$ 表示召回率从第 $k-1$ 个样本到第 k 个样本的变化量.

3.3 训练结果

改进前后的模型在经过 150 次 epoch 迭代后模型的性能评价曲线显示在图 14 中, 图 14 中曲线绘制使用了平滑处理. 图 14(a) 为改进的 YOLOv5-Upgraded 以及朴素 YOLOv5s 的训练损失, 损失函数可以直观地反映网络模型能否随着迭代次数的增加而稳定收敛. 随着迭代次数的增加, 两者均平缓下降并趋于平稳, 与朴素 YOLOv5s 相比, YOLOv5-Upgraded 的训练损失更小, 收敛更快, 证明 YOLOv5-Upgraded 对闪电哨声波的特征识别效果较强.

mAP 用于评估模型的性能, 图 14(b) 表明在经过 150 次迭代后, YOLOv5-Upgraded 的 mAP 达到 99.09%, 并趋于稳定; 而朴素 YOLOv5s 的 mAP 为 96.4%, 并且在第 80 次迭代后逐渐降低, 这说明 YOLOv5-Upgraded 具有较强的泛化能力, 模型性能更加稳定.

精度与召回率同样可以用来评估模型的性能, 精度表明检测出的闪电哨声波正确的比例, 精度越大, 证明模型错误检测出闪电哨声波的情况越少, 即越容

易识别闪电哨声波特征; 召回率表明检测出闪电哨声波的能力, 召回率越大, 表明模型越不容易漏检闪电哨声波. 图 14(c)(d) 分别给出的是模型改进前后的精度以及召回率曲线, 可以看出在经过 150 次迭代后 YOLOv5-Upgraded 精度和召回率趋于平稳, 分别达到 98.82% 以及 96.20%, 而朴素 YOLOv5s 的精度低于改进后模型, 并且在第 80 次迭代后召回率逐渐降低, 表明 YOLOv5-Upgraded 更容易提取闪电哨声波的特征进而正确识别, 综合性能优于朴素 YOLOv5s.

3.4 改进模型与基础网络性能对比实验

为了充分评估改进后模型的性能, 实验使用相同数据集, 在多个模型上进行对比实验, 以平均精度均值 (mAP)、召回率 (Recall)、模型权重 (Size)、检测速度 (FPS) 作为评价指标, 记录 150 次迭代后各个模型的实验结果, 如表 4 所示. 可以看出, YOLOv5-Upgraded 的平均精度为 99.09%, 相比 YOLOv3, YOLOv5s, YOLOv5m, YOLOv5l, 分别提高了 5.89%, 2.75%, 2.40%, 3.71%, 平均精度最高. 从召回率来看, YOLOv5-Upgraded 最高, 比 YOLOv3 和 YOLOv5s 分别提高了 5.07% 和 9.62%, 表明其检测

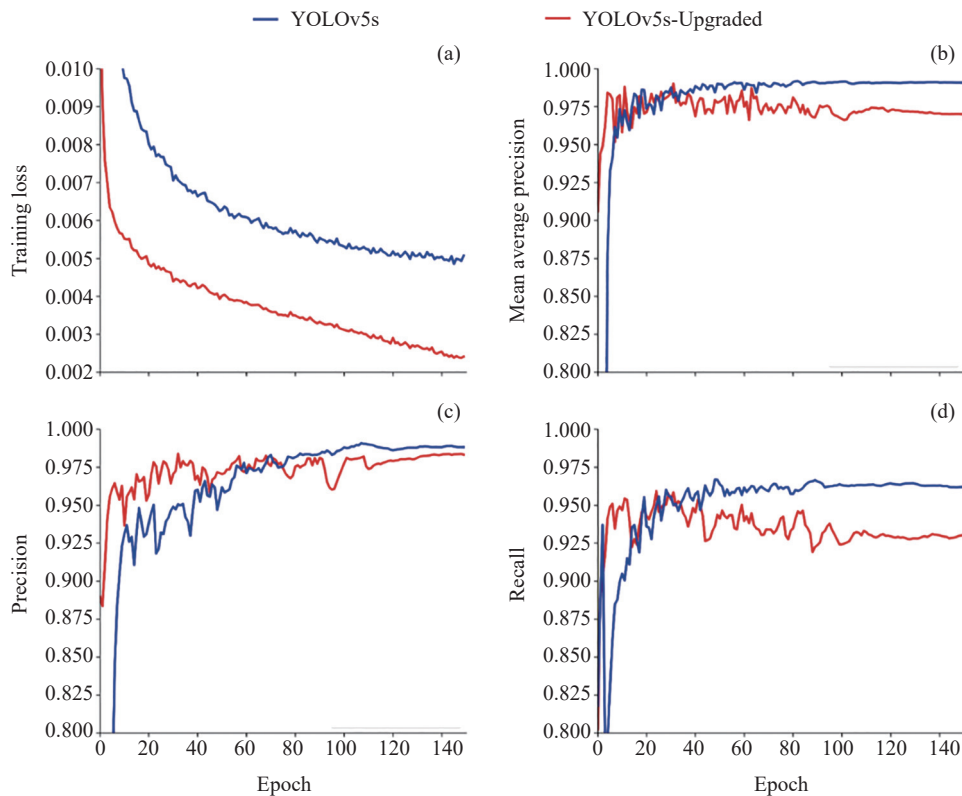


图 14 模型训练结果

Fig. 14 Network training results

表 4 五种目标检测网络性能比较
Table 4 Performance comparison of five object detection networks

模型	平均精度	召回率/(%)	Size/(MB)	FPS
YOLOv3YOLOv5s	93.2096.34	86.5891.07	117.7713.77	80105
YOLOv5m	96.69	93.21	40.08	103
YOLOv5 IYOLOv5-Upgraded	95.3899.09	89.1596.20	88.3613.78	7112

出闪电哨声波的能力优于其他算法,显著降低了闪电哨声波的漏检率.从模型大小看,YOLOv5-Upgraded的大小为13.78 MB,远远小于YOLOv3和YOLOv5m以及YOLOv5l,仅比YOLOv5s大了0.01 MB,但在平均精度和召回率上有大幅度提升.从FPS来看,YOLOv5-Upgraded中插入了CA注意力机制,加快了预测边框的回归速度,达到了112帧,高于其他模型.根据实验结果可以得出,YOLOv5-Upgraded在实现模型轻量化的同时,提高了闪电哨声波的检测精度以及降低了漏检率,综合性能最优.

3.5 不同注意力机制的性能对比实验

为了验证引入CA注意力机制可以帮助模型更准确提取到闪电哨声波的特征,本研究在损失函数改进为SIoU、激活函数改进为Mish的基础上插入不同注意力机制,例如SE,CBAM,ECA进行对比实验,从平均精度指标上给出了融合不同注意力机制的实验结果,如图15所示,图15曲线绘制使用了平滑处理.可以看出融合CA注意力机制的YOLOv5-Upgraded最高,这说明对于闪电哨声波的特征提取,不仅需要考虑到特征图通道之间的关系,还需要考虑到位置关系,使用CA注意力机制进行特征提取有更好的效果.

3.6 消融实验

为了进一步验证改进的各个模块对模型的影响,本研究进行5组消融实验,如表5和图16所示,YOLOv5s-SIoU,YOLOv5s-Mish,YOLOv5s-CA分别对应只更换损失函数为SIoU、只更换激活函数为Mish以及只插入CA注意力机制模块.其中因在多个epoch后网络参数未有更新,YOLOv5s-SIoU在第134个epoch结束训练,YOLOv5s-CA在第143个epoch后结束训练.

图16(a)为模型训练时使用了平滑处理后的平均精度变化曲线,在网络模型训练的前20个epoch内迅速上升,在第100个epoch后趋于稳定,YOLOv5-Upgraded达到了99.09%,仅比YOLOv5s-SIoU低0.11%,高于其他模型;图16(b)为模型训练时的召回

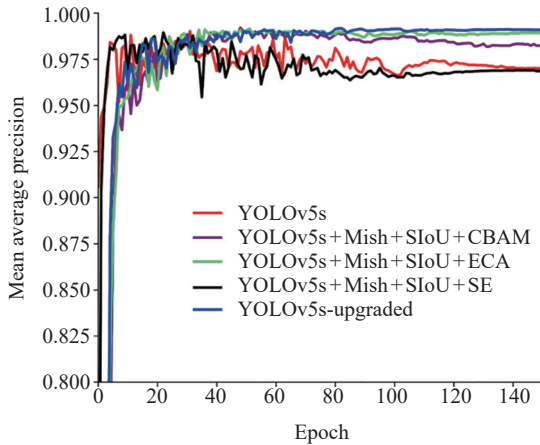


图 15 不同注意力机制的 mAP 对比实验

Fig. 15 mAP comparison experiment of different attention mecanisms

表 5 消融实验结果
Table 5 Ablation experimental results

模型	Mish	SIoU	CA	平均精度 (0.5)/(%)	召回率 (%)
YOLOv5s	×	×	×	96.34	90.83
YOLOv5s-Mish	✓	×	×	99.20	93.37
YOLOv5s-SIoU	×	✓	×	98.41	94.06
YOLOv5s-CA	×	×	✓	97.53	94.70
YOLOv5-Upgraded	✓	✓	✓	99.09	96.20

率变化曲线,在前20个epoch内迅速上升,在第120个epoch后逐渐平缓,最终YOLOv5-Upgraded的召回率达到96.20%,相比其他模型,有较大幅度提升.

3.7 测试集对比实验

使用YOLOv5s与YOLOv5-Upgraded分别在测试集上检测,图17和图18是为部分测试集上的结果,图中标号相同的表示检测的为同一张图,图中红框上的数字为置信度,表示模型认为红框内为闪电哨声波的概率.经对比可知,对于较为明显的闪电哨声

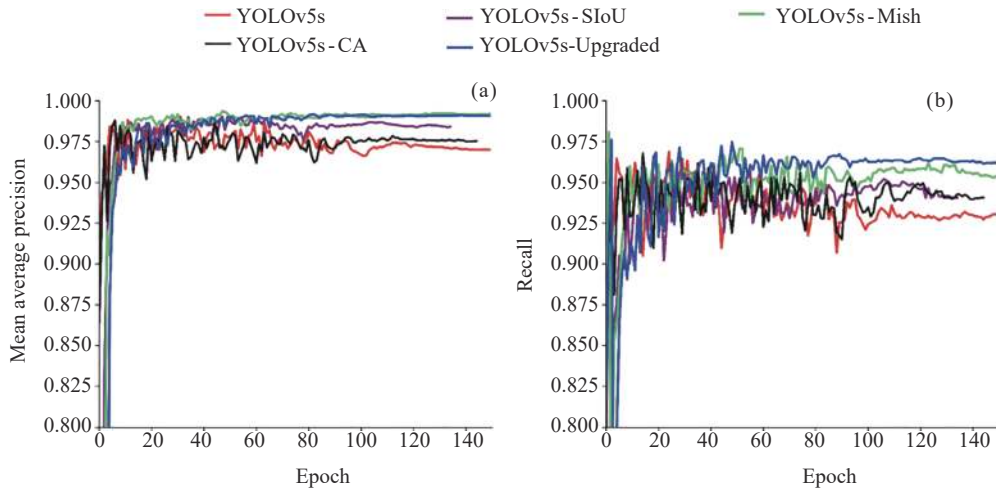


图 16 不同改进模型的结果

Fig. 16 Results of different improved models

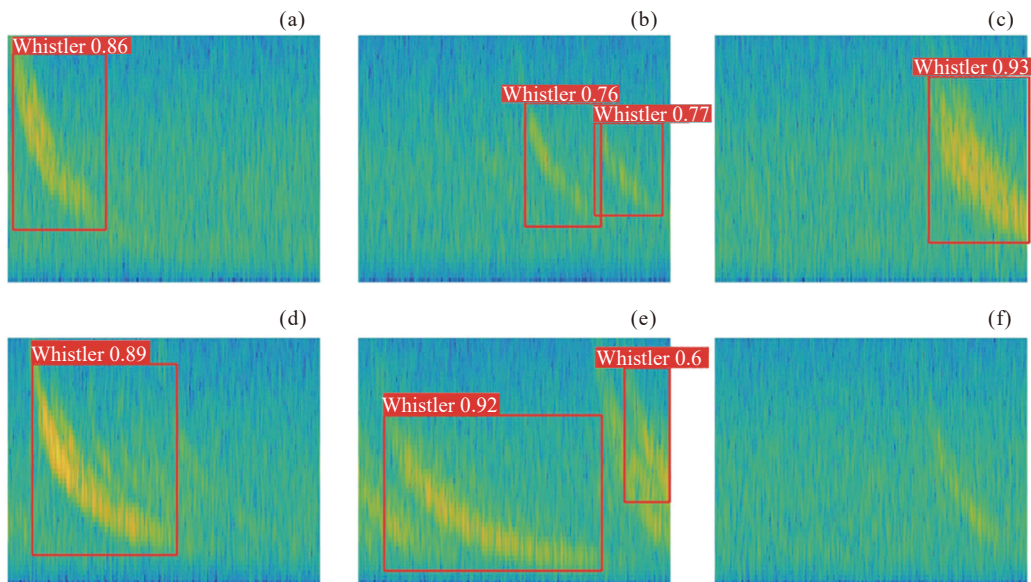


图 17 YOLOv5s 识别结果示例

Fig. 17 Examples of YOLOv5s identification results

波, 改进前后的模型都可以识别到, 但 YOLOv5-Upgraded 的置信度明显高于 YOLOv5s, 如图 17 和图 18 中 (a)~(d) 所示, 对于较为密集以及特征不明显的闪电哨声波, YOLOv5s 出现了漏检现象, 此时 YOLOv5-Upgraded 仍旧能够检测, 如图 17 和图 18 中的 (e)(f) 所示. 可以得出, 改进后的模型大大降低了闪电哨声波的漏检率, 极大地提高了模型的鲁棒性, 证明了模型的有效性.

3.8 轨道对比实验

为了进一步检验改进后模型的实际性能, 本研究

使用 YOLOv3, YOLOv5s, YOLOv5-Upgraded 三种模型在包含上百个闪电哨声波的实际轨道上进行检测, 并对检测情况进行分类总结, 得到以下结果. 对于 L 形态特征较为明显的闪电哨声波, 未出现两种情况: 一是 3 种模型都检测不到, 二是 YOLOv5s 检测到, YOLOv3 和 YOLOv5-Upgraded 检测不到.

如图 19 所示, 图 19 (a) 中闪电哨声波的 L 形态较为明显, 以上 3 种模型均能够识别到; 其中 YOLOv5-Upgraded 的置信度 84%, 为其他二种模型的 2 倍多. 图 19 (b)(d) 为 3 种不同形态的闪电哨声波, 3 个模型

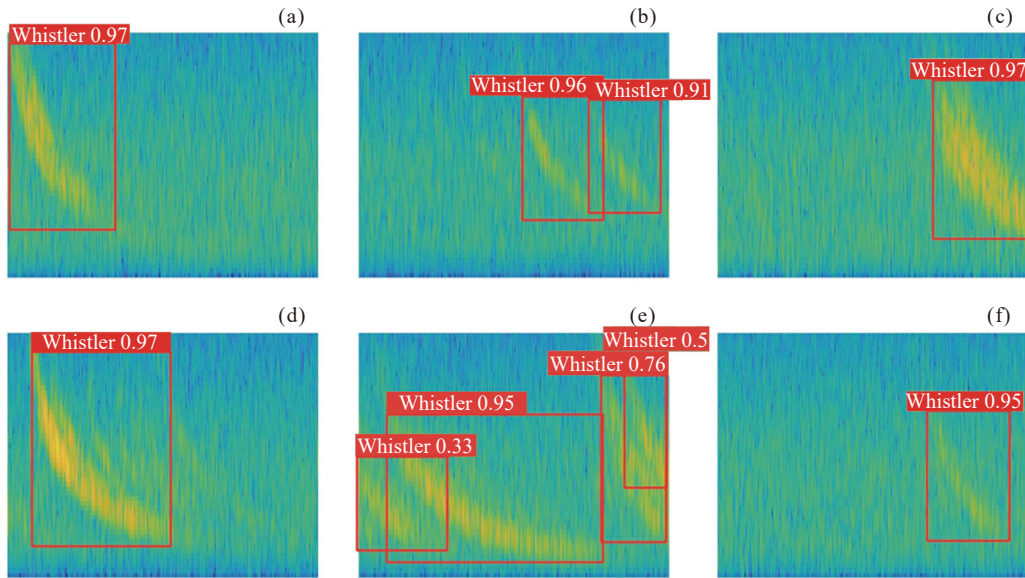


图 18 YOLOv5-upgraded 识别结果示例

Fig. 18 Examples of YOLOv5-upgraded identification results

中分别仅有 1 个模型识别不到, 另外 2 个可以识别到; 但是, 与图 19 (a) 相比, 单一检测模型的置信度较低, 该结果提示可构造集成的深度学习框架, 以进一步降低闪电哨声波的漏检率. 图 19 (e) 为只有 YOLOv5-Upgraded 一个模型检测到的情况, 从视觉上看图 19 (e) 目标的能谱密度较低且发散, 改进的模型以 92% 置信度检测出其他 2 个模型未检测出的闪电哨声波, 表明考虑通道与位置关系的 CA 注意力机制更易于提取该形态特征. 图 19 (f) 为只有 YOLOv3 一个模型检测到的情况, 该类情况仅在一个轨道上出现, 该轨道特征图每隔一秒出现明显的竖状噪声, 证明基于 YOLOv5s 改进的模型虽然在整体上大幅度降低了漏检率, 但由于模型深度远小于 YOLOv3, 在极端情况下如背景噪声明显时模型仍有出现漏检的情况, 下一步改进目标为进一步降低模型在复杂背景噪声环境下闪电哨声波的漏检率.

4 结语

针对现有闪电哨声波检测模型存在的漏检率高以及模型较大不适合部署到星载的问题, 本研究提出了一种基于 YOLOv5 的轻量化模型 YOLOv5-Upgraded. 改进主要体现在: 使用考虑预测边框与真实边框之间的方向的损失函数 SIoU 代替原有损失函数 CIoU, 使得模型更快地定位真实边框; 将 SiLU 激

活函数替换为具有更好梯度流的 Mish 激活函数, 其更加平滑的函数曲线带来更加稳定的梯度流, 使得信息更好地深入网络模型, 有效避免了模型训练过程中存在的梯度消失、梯度爆炸以及神经元坏死现象. 在主干网络中插入 CA 注意力机制, 降低了闪电哨声波的漏检率. 该模型与 YOLOv3 和 YOLOv5s 以及一些改进的网络进行对比实验, 得到以下结论.

基于改进 YOLOv5 的模型 YOLOv5-Upgraded, 平均精度均值 (mAP) 为 99.09%, 召回率 (Recall) 为 96.20%, 与朴素 YOLOv5s 相比, 分别提升了 2.75% 和 5.07%, 与 YOLOv3 相比, 分别提高了 5.89% 和 9.62%. YOLOv5-Upgraded 模型大小仅为 13.78 MB, 较占用 117.77 MB 的 YOLOv3 检测模型节约了 88.29% 的内存资源. 基于智能语音的 LSTM 闪电哨声波识别模型大小为 82.89 MB, 相比之下节约了 83.38% 的内存资源. 研究表明改进后的模型降低了闪电哨声波的漏检问题, 实现了模型轻量化, 并且易于部署到卫星设备, 极大地提高了星载识别的可能性, 有望为闪电定位以及雷电预警等相关应用提供数据支撑.

本文基于 YOLOv5 模型对闪电哨声波的检测进行研究, 虽然取得较好的检测效果, 但仍存在不足, 在之后的工作中会寻求改进.

(1) YOLOv5 为典型的 CNN 模型, 该类模型在捕获图像局部信息时有较好效果, 但对于捕获图像全

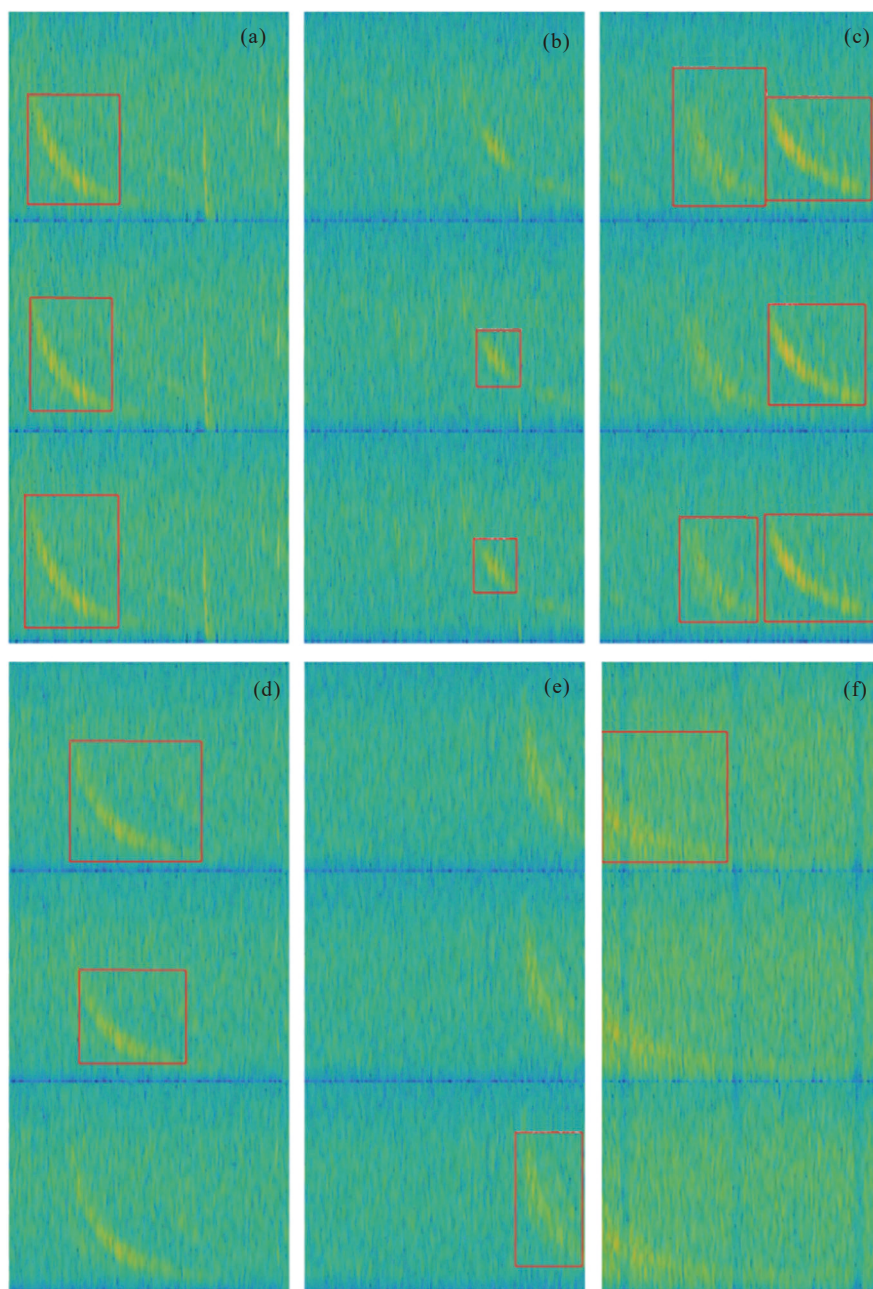


图 19 三种模型在轨道上的对比

Fig. 19 Comparison of three models on orbit

局信息时存在较大问题. 后续会在模型中引入采用自注意力机制的 Transformer 模块来增强模型捕获图像全局信息的能力.

(2) 本文提出的模型在检测大尺度闪电哨声波时性能较好, 但是数据集中小尺度闪电哨声波数量较少, 后续学术研究会采用增大数据集以及引入小目标数据增强策略, 以提升小尺度闪电哨声波的检测效果, 进而提高模型的健壮性.

参考文献

- [1] SANTOLÍK O, PARROT M, INAN U S, *et al.* Propagation of unducted whistlers from their source lightning: A case study[J]. *Journal of Geophysical Research: Space Physics*, 2009, **114**(A3): A03212
- [2] YUAN Jing, WANG Qiao, ZHANG Xuemin, *et al.* Advances in the automatic detection algorithms for lightning whistlers recorded by electromagnetic satellite data[J]. *Chinese Journal of Geophysics*, 2021, **64**(5): 1471-1495 (袁静, 王桥, 张学民, 等. 基于电磁卫星的闪电哨声波智能检测

- 算法的研究进展[J]. 地球物理学报, 2021, **64**(5): 1471-1495
- [3] BAYUPATI I P A, KASAHARA Y, GOTO Y. Study of dispersion of lightning whistlers observed by akebono satellite in the Earth's plasmasphere[J]. *IEICE Transactions on Communications*, 2012, **E95.B**(11): 3472-3479
- [4] PARROT M, BENOIST D, BERTHELIER J J, *et al.* The magnetic field experiment IMSC and its data processing onboard DEMETER: Scientific objectives, description and first results[J]. *Planetary and Space Science*, 2006, **54**(5): 441-455
- [5] ZEREN Zhima, LIU Dapeng, SUN Xiaoying, *et al.* Current status and scientific progress of the Zhangheng-1 satellite mission[J]. *Reviews of Geophysics and Planetary Physics*, 2023, **54**(4): 455-465 (泽仁志玛, 刘大鹏, 孙晓英, 等. 张衡一号电磁卫星在轨情况及主要的科学成果[J]. 地球与行星物理理论评(中英文), 2023, **54**(4): 455-465)
- [6] CAO J B, ZENG L, ZHAN F, *et al.* The electromagnetic wave experiment for CSES mission: Search coil magnetometer[J]. *Science China Technological Sciences*, 2018, **61**(5): 653-658
- [7] HUANG J P, LEI J G, LI S X, *et al.* The Electric Field Detector (EFD) onboard the ZH-1 satellite and first observational results[J]. *Earth and Planetary Physics*, 2018, **2**(6): 469-478
- [8] ZEREN Z M, HUANG J P, SHEN X H, *et al.* Simultaneous observations of ELF/VLF rising-tone quasiperiodic waves and energetic electron precipitations in the high-latitude upper ionosphere[J]. *Journal of Geophysical Research: Space Physics*, 2020, **125**(5): e2019JA027574
- [9] HU Y P, ZEREN Z M, FU H S, *et al.* A large-scale magnetospheric line radiation event in the upper ionosphere recorded by the China-Seismo-Electromagnetic satellite[J]. *Journal of Geophysical Research: Space Physics*, 2023, **128**(2): e2022JA030743
- [10] WANG Y L, ZEREN Z M, WANG X, *et al.* Statistical characteristics of the local proton cyclotron band emissions observed by the CSES[J]. *Journal of Geophysical Research: Space Physics*, 2022, **127**(11): e2022JA030860
- [11] FISER J, CHUM J, DIENDORFER G, *et al.* Whistler intensities above thunderstorms[J]. *Annales Geophysicae*, 2010, **28**(1): 37-46
- [12] Oike Y, Kasahara Y, Goto Y. Spatial distribution and temporal variations of occurrence frequency of lightning whistlers observed by VLF/WBA onboard Akebono[J]. *Radio Science*: 12
- [13] ALI AHMAD U, KASAHARA Y, MATSUDA S, *et al.* Automatic detection of lightning whistlers observed by the plasma wave experiment onboard the arase satellite using the OpenCV library[J]. *Remote Sensing*, 2019, **11**(15): 1785
- [14] KONAN O J E Y, MISHRA A K, LOTZ S. Machine learning techniques to detect and characterise whistler radio waves[OL]. arXiv preprint arXiv: 2002.01244, 2020
- [15] YUAN Jing, WANG Qiao, YANG Dehe, *et al.* Automatic recognition algorithm of lightning whistlers observed by the Search Coil Magnetometer onboard the Zhangheng-1 Satellite[J]. *Chinese Journal of Geophysics*, 2021, **64**(11): 3905-3924 (袁静, 王桥, 杨德贺, 等. 张衡一号感应磁力仪数据闪电哨声波自动识别[J]. 地球物理学报, 2021, **64**(11): 3905-3924)
- [16] YUAN Jing, WANG Zijie, ZEREN Zhima, *et al.* Automatic recognition algorithm of the lightning whistler waves by using speech processing technology[J]. *Chinese Journal of Geophysics*, 2022, **65**(3): 882-897 (袁静, 王子杰, 泽仁志玛, 等. 基于智能语音技术的闪电哨声波自动识别[J]. 地球物理学报, 2022, **65**(3): 882-897)
- [17] ZEREN Zhima, SHEN Xuhui, CAO Jinbin, *et al.* Statistical analysis of ELF/VLF magnetic field disturbances before major earthquakes[J]. *Chinese Journal of Geophysics*, 2012, **55**(11): 3699-3708 (泽仁志玛, 申旭辉, 曹晋滨, 等. 强震前 ELF/VLF 磁场的扰动特征统计研究[J]. 地球物理学报, 2012, **55**(11): 3699-3708)
- [18] HU Yunpeng, ZEREN Zhima, HUANG Jianping, *et al.* Algorithms and implementation of wave vector analysis tool for the electromagnetic waves recorded by the CSES satellite[J]. *Chinese Journal of Geophysics*, 2020, **63**(5): 1751-1765 (胡云鹏, 泽仁志玛, 黄建平, 等. 张衡一号卫星记录的空间电磁波传播特征分析方法及算法实现[J]. 地球物理学报, 2020, **63**(5): 1751-1765)
- [19] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016
- [20] MISRA D. Mish: A Self Regularized Non-Monotonic Activation Function[M/OL]. arXiv, 2020
- [21] GEVORGYAN Z. Siou loss: More powerful learning for bounding box regression[OL]. arXiv preprint arXiv: 2205.12740, 2022
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018
- [23] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 13708-13717

作者简介



路超男, 2000年2月出生, 现在中国科学院大学攻读硕士学位, 专业为计算机技术。

E-mail: 1104193785@qq.com



泽仁志玛(通信作者)女, 1976年出生, 现为应急管理部国家灾研院研究员, 主要从事电磁卫星数据处理及科学应用相关工作。

E-mail: zerenzhima@qq.com