

HAN Jinsheng, YUAN Jing, WANG Qiao, LIU Qinqin, YIN Hanke, LIU Haijun, ZHAO Shufan, SHEN Xuhui, WANG Yali. Fast and Robust Automatic Extraction Method for the Lightning Whistler Scattering Coefficient of the Zhangheng Satellite (in Chinese). *Chinese Journal of Space Science*, 2025, 45(4): 960–974. CSTR:32142.14.cjss.2023–0127. DOI:10.11728/cjss2025.04.2023–0127

快速稳健的张衡卫星闪电哨声波散射系数自动提取方法*

韩金昇¹ 袁 静¹ 王 桥² 刘芹芹³ 银韩柯¹
刘海军¹ 赵庶凡³ 申旭辉³ 王亚丽¹

1(防灾科技学院信息工程学院 廊坊 065201)

2(中国地震局地球物理研究所 北京 100081)

3(中国科学院国家空间科学中心 北京 100190)

摘 要 张衡卫星日产数据高达 20 GByte, 人工方法已经无法满足海量数据提取的需求. 提出了一种快速稳健的闪电哨声波散射系数自动提取方法. 将原始数据转换为时频图和音频文件; 针对时频图, 创建 YOLOV5 神经网络自动定位闪电哨声波, 输出其时频位置信息; 根据其时频位置信息截取对应音频文件得到含有闪电哨声波的音频数据, 并用零将其填充为 0.8 s 的音频片段; 提取音频片段的梅尔频率倒谱系数, 将其输入多头注意力机制改进的门控循环网络, 自动提取闪电哨声波散射系数. 将该方法应用于 2020 年 2 月的数据集, 得到如下结果: 其平均绝对误差和平均绝对百分比误差为 0.453 和 0.176, 相比文献 [1] 方法降低了 70% 和 46%. 每段数据平均处理时间为 0.074 s, 相比文献 [1] 方法降低了 92%. 本文提出的自动音频分析方法, 可较为准确和快速地提取闪电哨声波的散射系数.

关键词 张衡卫星, GRU 神经网络, YOLOV5 神经网络, 闪电哨声波, 散射系数

中图分类号 P352

Fast and Robust Automatic Extraction Method for the Lightning Whistler Scattering Coefficient of the Zhangheng Satellite

HAN Jinsheng¹ YUAN Jing¹ WANG Qiao² LIU Qinqin³ YIN Hanke¹
LIU Haijun¹ ZHAO Shufan³ SHEN Xuhui³ WANG Yali¹

1(School of Information Engineering, Institute of Disaster Prevention, Langfang 065201)

2(Institute of Geophysics, China Earthquake Administration, Beijing 100081)

* 国家自然科学基金青年基金项目 (42104159) 和中国地震局教师基金项目 (20150109) 共同资助

2023-11-13 收到原稿, 2024-07-08 收到修定稿

©The Author(s) 2025. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

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

Abstract The daily production data of the Zhangheng satellite can reach up to 20 GB, rendering manual methods inadequate for handling such massive data demands. This paper proposes a rapid and robust Lightning Whistle Scattering Coefficient Automatic Extraction Method (LWSC-AEM). Firstly, detailed data from the Search Coil Magnetometer (SCM) of the Zhangheng satellite is extracted using a sliding window of 0.8 seconds, which is then transformed into time-frequency plots and audio files. Secondly, a YOLOV5 neural network is employed to automatically locate LW in the time-frequency plots and output their time-frequency position information. Subsequently, the corresponding audio data containing Lightning Whistlers is extracted based on this time-frequency position information from the files, and zero-padded to form audio segments of 0.8 seconds. Finally, the Mel Frequency Cepstral Coefficients (MFCCs) of the audio segments are extracted and fed into a Gate Recurrent Unit (GRU) improved with a multi-head attention mechanism to automatically extract the LW scattering coefficient. Applying this method to the data from the VLF band of the SCM payload of the Zhangheng satellite in February 2020 yields the following results: the average absolute error and average absolute percentage error are 0.453 and 0.176 respectively. Compared to the method by Ref. [1], the average absolute error is reduced by 1.079, a decrease of 70%, and the average absolute percentage error is reduced by 0.148, a decrease of 46%. The average processing time per data segment is 0.074 seconds, which is a reduction of 0.826 seconds, or 92%, compared to the method by Ref. [1], which processed each data segment on average. The automatic extraction method for lightning whistler wave scattering coefficients proposed in this paper can quickly and accurately extract these coefficients.

Key words Zhangheng satellite, GRU neural network, YOLOV5 neural network, Lightning whistler, Scattering coefficient

0 引言

闪电激发的哨声波以近似纵向的右旋波模式穿透电离层, 通常沿地球磁力线在两个半球的磁共轭点之间传播, 其传播路径经过电离层和磁层, 这使得其携带了大量的地球空间环境信息. 因此, 闪电哨声波 (Lightning Whistler) 成为探测空间物理环境的重要媒介^[2-10]. 例如, 借助磁尾的闪电哨声波探测等离子体波在空间等离子体中的能量转换过程^[11]; 借助磁层顶的闪电哨声波探测磁重联过程中闪电哨声波的产生机制及其对电子动力学的影响^[12].

闪电哨声波的观测源于 20 世纪 80 年代, 其主要借助散点状的地面探测技术观测闪电哨声波^[13,14], 空间局限性大; 随着卫星观测技术的迅速发展, 全球闪电哨声波数据与日俱增. 以往有关闪电哨声波的分析, 主要通过人工手动分析地面雷达或卫星观测到的电磁波信号来获取其物理参数^[3,15,16], 庞大的工作量导致所获得的有效信息十分有限^[2,13,17], 这使得人类未能掌握全球闪电哨声波传播特征和介质参数的时

空分布. 2008 年 Lichtenberger 等^[18]率先提出, 依赖人工处理闪电哨声波已成为探测空间物理环境的瓶颈. 随后开展了一系列的闪电哨声波智能处理方面的工作^[5,19-24], 这些工作主要聚焦于闪电哨声波事件的自动识别, 无法自动获取闪电哨声波的物理参数, 例如散射系数. 闪电哨声波的散射系数最早由 Eckersley^[25]发现, 其描述了闪电哨声波的频率与时间之间的关系, 即

$$t = D/\sqrt{f} + t_0. \quad (1)$$

同时反映了在给定的频率和方向下, 单位体积内的电子对电磁波的散射强度. 式 (1) 中, D 为散射系数, (t, f) 为闪电哨声波轨迹的时频点对 (时间, 频率), t_0 为闪电触发初始时间. Bayupati 等^[6]进一步指出散射系数 D 是开展闪电哨声波反演地球等离子体层中电子密度分布的关键物理参数.

文献 [1] 首次借助人工智能技术探索闪电哨声波散射系数自动提取方法 (Unet-WH), 该方法以闪电哨声波时频图为研究对象, 利用 Unet 神经网络和图

像处理方法自动提取闪电哨声波散射系数,取得了一定效果,但仍存在以下缺点.

(1)Unet-WH 方法依赖人工设置阈值致使算法的鲁棒性不足.如图 1(a)所示,该方法对分散且连通性较强的闪电哨声波散射系数提取效果好;如图 1(b)(c)所示,其对密集或连通性弱的闪电哨声波散射系数提取效果差.分析原因可以看出,使用 Unet 神经网络对闪电哨声波进行描述后,每条闪电哨声波轨迹中会出现缝隙(连通性弱的闪电哨声波轨迹中出现的缝隙大,连通性强的闪电哨声波轨迹中出现的缝隙小),导致闪电哨声波轨迹变为多个连通域,而 Unet 神经网络并不知道某一条闪电哨声波轨迹是由哪些连通域组成的,所以进行图像腐蚀来减小缝隙并合并连通域,设置的图像腐蚀的程度需要人工设置阈值,如果图像腐蚀的程度过大,则可能会出现相邻的闪电哨声

波的连通域连在一起的情况.进行图像腐蚀之后,还会存在属于一条闪电哨声波轨迹的但连通域没有连在一起的情况,所以进行连通域合并操作将连通域合并在一起,设置连通域合并距离的大小需要人工设置阈值,但是闪电哨声波的持续时间和分布情况是不可知的,阈值过小,可能将一条哨声波分为多个连通域;阈值过大,则可能将多条哨声波的连通域连在一起,所以使用人工设置阈值提取散射系数方法鲁棒性差.

(2)Unet-WH 方法速度慢,原因是其采用的 Unet 神经网络模型计算成本高,计算效率低.

为了克服上述问题,本文提出了快速稳健的闪电哨声波散射系数自动提取方法.利用 YOLOV5 神经网络对闪电哨声波进行识别,进而利用语音和多头注意力机制改进的 GRU 神经网络自动提取散射系数.

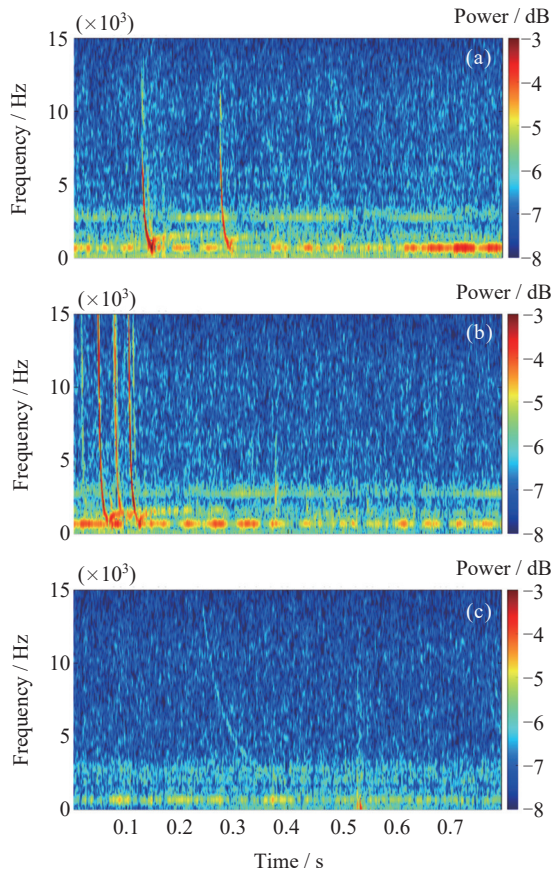


图 1 闪电哨声波. (a) 分散且连通性好的哨声波, (b) 密集哨声波, (c) 连通性差的哨声波

Fig. 1 Lightning Whistler. (a) Well-spaced and well-connected, (b) tightly spaced, (c) poor connectivity

1 自动提取算法

1.1 算法整体框架

算法框架如图 2 所示,主要分为三个部分:预处理、闪电哨声波识别以及闪电哨声波散射系数的自动提取.

(1) 预处理阶段:通过 0.8 s 的滑动窗截取详查数据,将其转换为时频图 T 和音频文件 WAV.

(2) 闪电哨声波识别阶段:将时频图输入 YOLOV5 神经网络,识别单个闪电哨声波及其时频位置信息.

(3) 闪电哨声波散射系数自动提取阶段:根据闪电哨声波的时频位置信息从音频数据 WAV 中截取含有闪电哨声波的音频片段,并将其填充为 0.8 s 的音频片段 wav;提取音频片段 wav 的梅尔频率倒谱系数(MFCCs);将 MFCCs 特征输入到多头注意力机制改进的 GRU 神经网络 (GRU-MHA),输出闪电哨声波散射系数.

1.2 算法模型与数据提取

1.2.1 闪电哨声波识别模型

旨在使用 YOLOV5 神经网络在时频图上完成闪电哨声波事件的识别并输出其存续的时间范围. YOLOV5 神经网络模型由主干网络 (Backbone)、颈部网络 (Neck) 和头部网络 (Head) 组成,如图 3 所示.

主干网络:图 3 所示的 Backbone 部分,旨在提

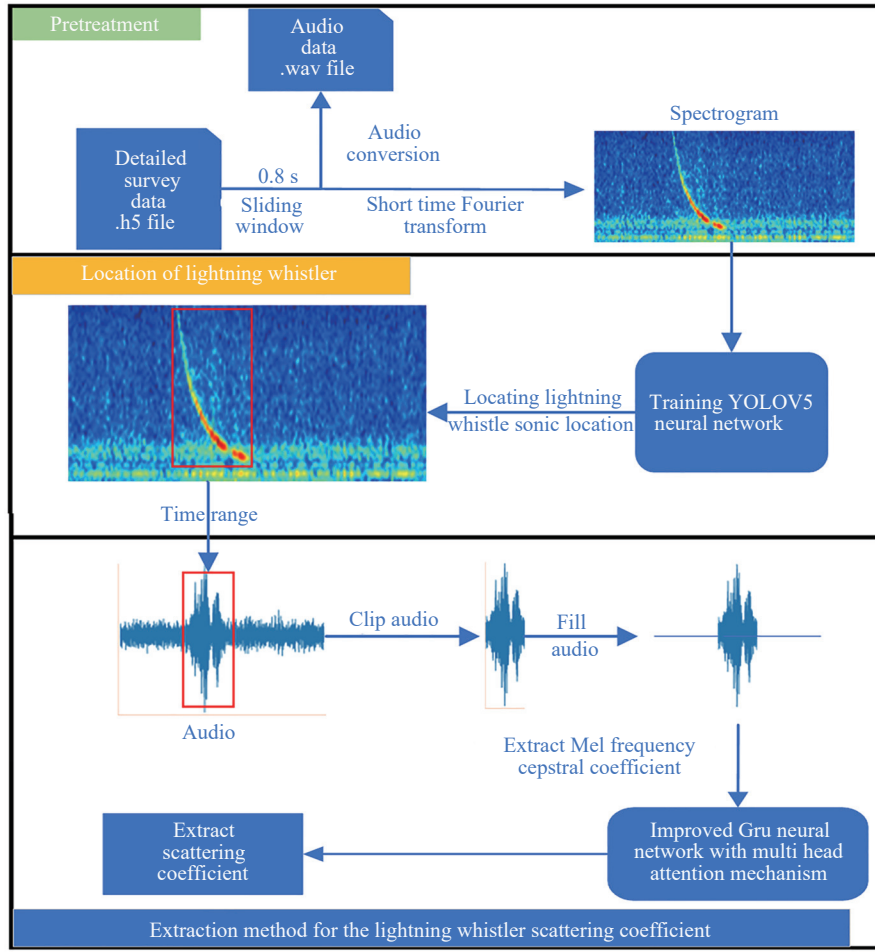


图2 算法框架

Fig. 2 Algorithm framework diagram

取和增强图像特征, 包括 CBS 层、C3 层和 SPPF 层. CBS 层用于特征提取; C3 层主要提取图像的语义信息 (Bottleneck_1 层采用残差结构, 而 Bottleneck_2 层采用常规卷积网络结构); SPPF 层解决卷积神经网络在提取重复特征时的问题, 加快候选框生成速度, 节省计算成本.

颈部网络: 图 3 所示的 Neck 部分, 旨在通过生成特征金字塔来增强模型对不同尺度闪电哨声波的检测, 从而精确识别不同尺度的闪电哨声波事件. Neck 部分采用 PANet 结构设计, 包括 FPN 结构和 PAN 结构. 其从多尺度特征金字塔中提取长距离和短距离的特征关系, 融合不同尺度的特征图, 丰富语义信息. FPN 结构自顶向下传递特征并通过上采样融合信息, PAN 结构为自底向上的特征金字塔, 通过下采样增强网络特征融合的能力.

头部网络: 图 3 所示的 Head 部分, 旨在从主干

网络或颈部网络获得的特征图中推断闪电哨声波的位置信息, 生成最终输出向量, 其中包括边界框、对象置信度得分和类别概率.

YOLOV5 的总损失函数 (L_{total}) 是三个损失函数的加权和, 其中包括闪电哨声波边界框损失 (L_{box})、闪电哨声波置信度损失 (L_{obj}) 和闪电哨声波分类损失 (L_{cls}), 有

$$L_{total} = \sum_{i=1}^N (\lambda_1 L_{box} + \lambda_2 L_{obj} + \lambda_3 L_{cls}). \quad (2)$$

其中, N 为样本数量, 即图像中存在的真实闪电哨声波边界框数量. λ_1, λ_2 和 λ_3 为损失函数的权重因子. L_{box} 用于衡量预测闪电哨声波边界框位置与真实闪电哨声波边界框之间的差异, L_{obj} 则用于评估预测框是否包含闪电哨声波以及置信度的准确性, L_{cls} 则用于衡量模型对闪电哨声波类别分类的准确性. 更详细的

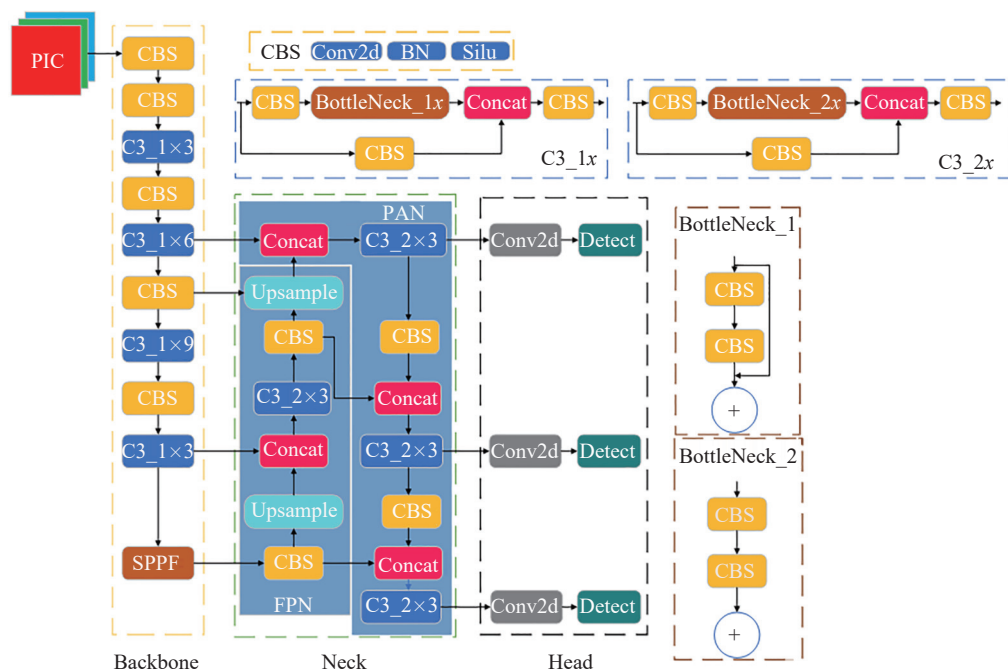


图3 YOLOV5 框架

Fig. 3 YOLOV5 framework diagram

细节请见参考文献 [26–28]. YOLOV5 神经网络识别效果如图 4 所示.

1.2.2 闪电哨声波散射系数自动提取

根据闪电哨声波识别结果获取含有闪电哨声波的音频, 提取其 MFCCs 特征; 进而构建 GRU 神经网络并添加多头注意力机制, 以实现闪电哨声波散射系数的自动提取.

获取哨声波音频. 根据闪电哨声波事件的识别结果, 如图 5(a) 中矩形框所示, 截取该时间范围的音频片段, 然后用零进行填充以获得长度为 0.8 s 的闪电哨声波音频数据.

提取 MFCC 特征. 闪电哨声波是时序信号, 通过语音播放器能够听到不同散射系数的“哨声”, 其 MFCCs 特征模拟了人耳的听觉机理, 具有显著的音频率列特征建模能力. 因此, 本文通过对音频片段进行预加重、分帧、加窗、傅里叶变换、梅尔滤波器组、对数运算和倒谱系数七个步骤提取闪电哨声波的 MFCCs 特征, 详细的提取过程请见参考文献 [24, 29, 30], 提取到的特征如图 5(b) 所示. 每一行代表一个时间帧, 共 96 个时间帧; 其每一列代表一维 MFCC 系数, 共 39 维 MFCC 系数 (13 维 MFCC 倒谱系数、13 维 MFCC 倒谱系数一阶差分系数和

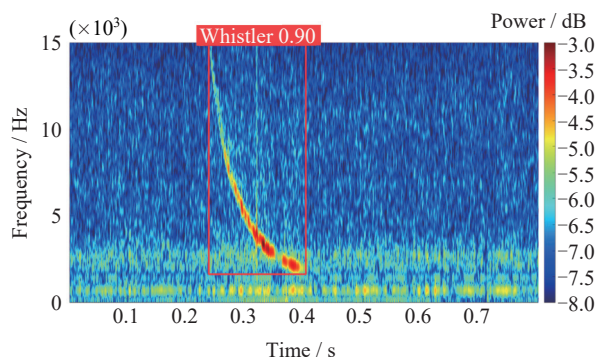


图4 YOLOV5 识别效果

Fig. 4 YOLOV5 identification result

13 维 MFCC 倒谱系数二阶差分系数), 闪电哨声波特征主要体现在水平方向 (用零填充音频导致闪电哨声波之外的特征是相同的).

构建多头注意力机制改进的 GRU 神经网络. 闪电哨声波的 MFCCs 特征具有明显的序列变化特点, 为了捕捉该信息, 采用循环结构简单的门控循环单元 (GRU) 神经网络, 同时考虑到闪电哨声波持续时间短, 其变化特征在整个 MFCCs 特征中所占的比例小. 为了增强 GRU 神经网络对不同散射系数的闪电哨声波的 MFCCs 变化信息的提取能力, 采用多头注意力机制改进 GRU 神经网络 (见图 6), 其主要由四部分

组成: 输入层、GRU 层、多头注意力机制层和输出层。

输入层. 输入层由 $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \dots, \mathbf{X}_t$ 组成, 是 t 个闪电哨声波的 MFCCs 特征, 单个闪电哨声波的 MFCCs 特征是维度为 96×39 , 即输入层是维度为 $t \times 96 \times 39$ 。

GRU 层. GRU 层是由若干个 GRU 神经元组成的神经网络, GRU 神经元结构如图 7 所示. 其完整的计算过程如下:

$$\mathbf{r}_t = \sigma(\mathbf{W}_r^h \mathbf{h}_{t-1} + \mathbf{W}_r^x \mathbf{x}_t + \mathbf{b}_r), \quad (3)$$

$$\mathbf{z}_t = \sigma(\mathbf{W}_z^h \mathbf{h}_{t-1} + \mathbf{W}_z^x \mathbf{x}_t + \mathbf{b}_z), \quad (4)$$

$$\mathbf{h}_t = \mathbf{z}_t \odot \mathbf{h}_{t-1} + (1 - \mathbf{z}_t) \odot \tilde{\mathbf{h}}_t, \quad (5)$$

$$\tilde{\mathbf{h}}_t = \tanh[\mathbf{W}_h^h(\mathbf{r}_t \odot \mathbf{h}_{t-1}) + \mathbf{W}_h^x \mathbf{x}_t + \mathbf{b}_h]. \quad (6)$$

其中, $\sigma(\cdot)$ 为 sigmoid 激活函数; $\mathbf{W}_r^h, \mathbf{W}_r^x, \mathbf{b}_r, \mathbf{W}_z^h, \mathbf{W}_z^x, \mathbf{b}_z$ 为可训练的网络权重参数; $\odot$ 为哈达玛乘积; $\mathbf{h}_{t-1}$ 表示输入上一个神经元中闪电哨声波的 MFCCs 特征的隐藏信息; $\mathbf{x}_t$ 表示输入当前神经元中

的闪电哨声波的 MFCCs 特征信息; $\mathbf{r}_t$ 为重置门, 用来控制 $\mathbf{h}_{t-1}$ 中信息的忽略程度, 重置门越大, $\mathbf{h}_{t-1}$ 被清除的信息越多, 留下的信息越少; $\mathbf{z}_t$ 为更新门, 用于控制 $\mathbf{h}_{t-1}$ 中的信息的保留程度, 更新门的值越大, 说明 $\mathbf{h}_{t-1}$ 中信息保留的越多; $\tilde{\mathbf{h}}_t$ 是神经元的候选隐藏信息, 由两部分组成, 一部分由重置门和 $\mathbf{h}_{t-1}$ 决定, 另一部分是输入当前神经元的闪电哨声波的 MFCCs 特征信息; $\mathbf{h}_t$ 为神经元的隐藏信息, 会被输入到下一个神经元中, 其由两部分组成, 一部分由 $\mathbf{h}_{t-1}$ 和更新门决定, 另一部分由候选隐藏信息 $\tilde{\mathbf{h}}_t$ 和更新门决定. 对于损失函数, 闪电哨声波散射系数自动提取属于回归问题, 所以该网络使用 Huber Loss 损失函数进行训练^[31], 即

$$L_\delta(\hat{y}_i, y_i) = \begin{cases} \frac{1}{2}(\hat{y}_i - y_i)^2, & |\hat{y}_i - y_i| \leq \delta; \\ \delta \left(|\hat{y}_i - y_i| - \frac{1}{2}\delta \right), & |\hat{y}_i - y_i| > \delta. \end{cases} \quad (7)$$

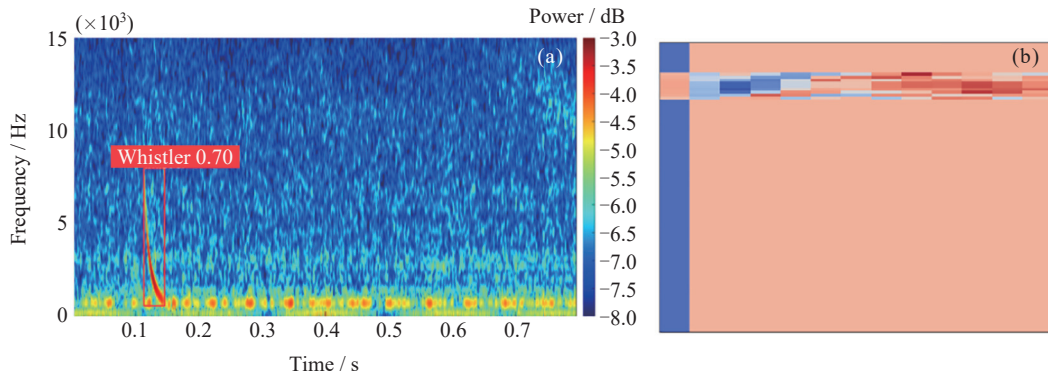


图 5 MFCCs 特征可视化. (a) 哨声波识别结果 (矩形框), (b) 哨声波的 MFCCs 特征

Fig. 5 Visualization of MFCCs features. (a) Whistler recognition results (shown in a rectangular box), (b) MFCCs characteristics of the whistler

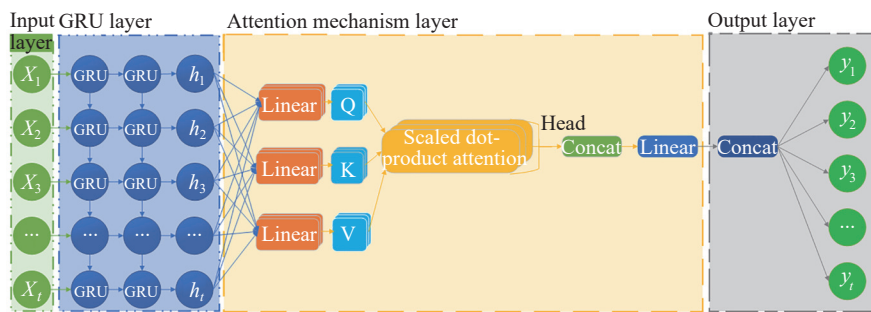


图 6 多头注意力机制改进的 GRU 网络

Fig. 6 GRU network with improved multi-head attention mechanism

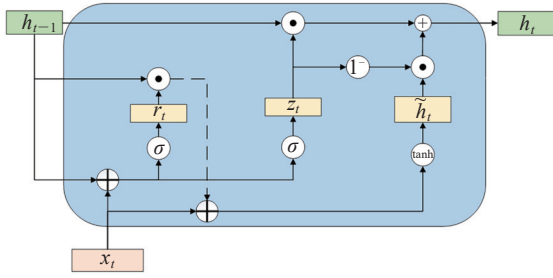


图7 GRU神经元的内部结构

Fig. 7 Internal structure diagram of GRU neurons

其中, $\hat{y}_i$ 为模型预测的散射系数, y_i 为真实的散射系数数值, δ 为损失函数中的一个超参数.

多头注意力机制层. 多头注意力机制层由四部分组成, 分别为线性层 (Linear)、注意力层 (Scaled dot-product attention)、拼接层 (Concat) 和全连接层 (Linear), 如图 8 所示. 具体如下.

线性层 (Linear): 创建三个权重矩阵分别为 W^Q , W^K 与 W^V ; 创建三个线性网络并将其从 GRU 层中得到的闪电哨声波的信息 h_t 分别与三个权重矩阵 W^Q , W^K 与 W^V 相乘得到 Query(Q), Key(K), Value(V), 计算过程如下:

$$\begin{cases} Q = h_t W^Q, \\ K = h_t W^K, \\ V = h_t W^V. \end{cases} \quad (8)$$

注意力层 (Scaled Dot-Product Attention): 注意力层旨在利用上一步得到的 Q , K , V 创建多个注意力 (head) 结构. 采用如下四个基本计算操作创建单个注意力 head 结构.

矩阵乘法操作 (Matmul):

$$S = \text{Matmul}(Q, K) = Q \cdot K^T. \quad (9)$$

缩放处理操作 (Scaled):

$$S' = S / \sqrt{d_k}. \quad (10)$$

激活操作 (Softmax):

$$S'' = \text{Softmax}(S') = \frac{e^{s'_i}}{\sum_{j=1}^C e^{s'_j}}. \quad (11)$$

矩阵乘法操作 (Matmul):

$$H = \text{Matmul}(S'', V) = S'' \cdot V^T. \quad (12)$$

其中, d_k 为 Q 和 K 的维度, C 为 $e^{s'_i}$ 的维度. 具体如图 9 所示. 将上述操作重复 n 次, 得到多头注意力结构 h_i , $i \in \{1, 2, \dots, n\}$.

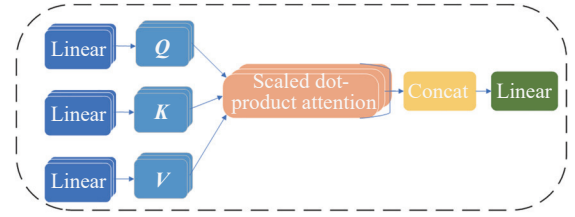


图8 多头注意力机制层结构

Fig. 8 Structure diagram of multi head attention mechanism layer

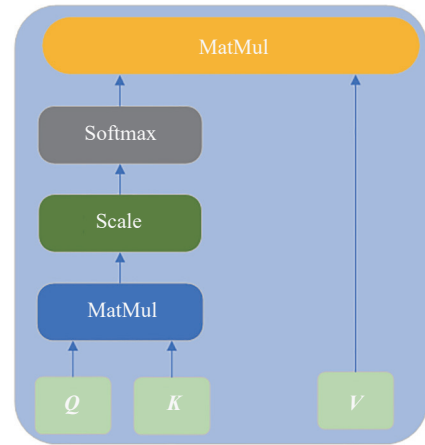


图9 缩放点积注意力

Fig. 9 Scaled dot-product attention

拼接层 (Concat): 将注意力层得到的多个注意力 h_i 进行拼接, 然后与额外的权重矩阵 W^o 相乘, 得到多头注意力. 拼接过程如下:

$$Z = [h_1, h_2, \dots, h_i]^T \cdot W^o. \quad (13)$$

全连接层 (Linear): 创建一个全连接网络, 将拼接层得到的多头注意力 Z 映射为一个含有所有注意力子空间特征信息的矩阵 L . 计算过程如下:

$$L = W_L Z + b_l. \quad (14)$$

其中, W_L 为全连接层中的权重参数, b_l 为全连接层的偏置向量.

输出层: 在输出层中, 将多头注意力机制产生的矩阵 L 输入到一个全连接网络, 由该网络输出最终的散射系数 y_t . 计算过程如下:

$$y_t = W_r L + b_r. \quad (15)$$

其中, W_r 为全连接网络的权重参数, b_r 为全连接层的偏置参数.

最终提取散射系数 D , 如图 10 所示. 图 10(a) 为

闪电哨声波轨迹, 图 10(b) 为人工提取和自动提取的闪电哨声波散射系数拟合的轨迹。紫线为人工提取的散射系数拟合轨迹, 黑线是自动提取的散射系数拟合轨迹。

2 实验结果与分析

为了验证所提出模型的性能, 对数据情况、评价指标、模型定量评价结果和模型定性评价结果进行分析。

2.1 数据分析

分析数据来自张衡一号卫星 2020 年 2 月 SCM 载荷 VLF 波段的详查数据, 以 0.8 s 的滑动窗截取详查数据, 对其进行短时傅里叶变换和音频转换得到时频图和音频文件。使用 0.8 s 的滑动窗截取详查数据的原因是, 在使用 YOLOV5 神经网络识别闪电哨声波和使用 GRU-MHA 网络提取闪电哨声波散射系数时, 都要保证闪电哨声波被完整地保存在一个时频图和音频文件中。如果闪电哨声波数据不完整, 将会严重影响 YOLOV5 神经网络识别效果和 GRU-MHA 模型提取效果。由于闪电哨声波持续时间长短不一, 统计张衡一号卫星 2020 年 2 月 25—29 日闪电哨声

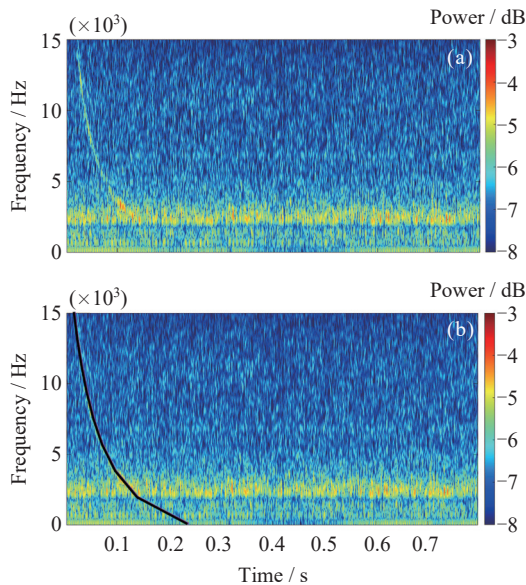


图 10 散射系数自动提取效果。(a) 含有闪电哨声波的时频图, (b) 散射系数提取结果

Fig. 10 Automatic extraction of scattering coefficient. (a) Time frequency plot containing lightning whistler, (b) scattering coefficient extraction result

波数据中持续时间小于 0.8 s 的闪电哨声波事件所占比例为 98.7%。所以选用了 0.8 s 的时间窗, 尽可能使闪电哨声波完整保存在一个时频图和音频文件中 (始终会存在被时间窗分割的闪电哨声波)。

本文提出的快速稳健的闪电哨声波散射系数自动提取方法 (闪电哨声波 SC-AEM) 主要由两个模型组成 (M1-YOLOV5 闪电哨声波自动识别模型和 M2-闪电哨声波散射系数自动提取模型), 为了评估模型性能, 首先选取 2 月 2 日到 2 月 24 日中的数据分别训练 M1 和 M2, 再采用 2020 年 2 月 25—29 日的数据对上述两个模型进行效果评价, 其训练集中含哨声波的数据大约为 6764 张时频图, 测试集共有时频图 73691 张及其相应的音频 73691 段。

2.2 评价指标

分别针对 YOLOV5 闪电哨声波自动识别模型和闪电哨声波散射系数自动提取模型两种不同模型制定不同的评价指标, 即自动识别效果评价指标和散射系数自动提取评价指标。

2.2.1 自动识别效果评价指标

精确度 (Precision, P). 目标检测网络的精确度

$$P = \frac{N_{tp}}{N_{tp} + N_{fp}}. \quad (16)$$

其中, N_{tp} 表示正确识别闪电哨声波事件的数量; N_{fp} 表示错误识别闪电哨声波的数量; 该指标值越大, 表明识别正确的数量越多, 也意味着识别错误的数量越少。

召回率 (Recall, R). 目标检测网络的召回率

$$R = \frac{N_{tp}}{N_{tp} + N_{fn}}. \quad (17)$$

其中, N_{tp} 表示正确识别闪电哨声波事件的数量; N_{fn} 表示漏检的闪电哨声波的数量。该指标值越大, 说明正确识别的闪电哨声波数量越多, 也意味着漏检的闪电哨声波越少。

2.2.2 散射系数自动提取评价指标

E_{MAE} 为平均绝对误差 (Mean Absolute Error, MAE), 散射系数的平均绝对误差

$$E_{MAE} = \frac{1}{N} \sum_{i=1}^N |N_A - D_0|. \quad (18)$$

其中, N_A 为人工提取的散射系数, D_0 为自动提取的散射系数, N 为闪电哨声波的数量。该指标值越小, 说明提取散射系数的效果越好, 与人工提取的散射系数

相差越小.

E_{MAPE} 为平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE): 散射系数的平均绝对百分比误差

$$E_{\text{MAPE}} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{D_a - D_0}{D_a} \right|. \quad (19)$$

其中, D_a 为人工提取的散射系数, D_0 为自动提取的散射系数, N 为闪电哨声波的数量. 该指标值越大, 说明提取散射系数的效果越差, 与人工提取的散射系数相差越大.

T_{AET} 为平均提取时间 (Average Extraction Time, AET), 散射系数的平均提取时间

$$T_{\text{AET}} = \frac{T}{N}. \quad (20)$$

其中, T 为提取散射系数的总时间, N 为闪电哨声波的数量.

2.3 模型定量评价

为了评估提出的闪电哨声波散射系数自动提取方法 (闪电哨声波 SC-AEM), 基于 2020 年 2 月 25—29 日的数据进行实验.

2.3.1 闪电哨声波识别

YOLOV5 神经网络识别闪电哨声波结果如表 1 所示, YOLOV5 神经网络识别闪电哨声波精确度为 91.8%, 召回率为 93.4%, 说明 YOLOV5 神经网络识别闪电哨声波效果很好, 识别错误和漏识别的情况很少, 这也保证了后续 GRU-MHA 提取散射系数结果的有效性. 相较于文献 [1], 使用 YOLOV5 对闪电哨声波进行识别准确率高, 每张时频图处理时间为 0.07 s. 文献 [1] 方法直接在时频图上提取散射系数效果差, 且每张时频图的处理时间为 0.9 s.

YOLOV5 神经网络漏识别的情况主要为将多条重叠的闪电哨声波事件识别为一个闪电哨声波事件, 如图 11 所示.

YOLOV5 神经网络识别错误的情况主要是将噪声误识别为闪电哨声波, 虽然其在形态上与闪电哨声波相似, 但音频中并无哨声, 具体情况如图 12 所示.

表 1 闪电哨声波识别结果

Table 1 Lightning whistle localization outcome

算法名称	$P/(%)$	$R/(%)$
YOLOV5	91.8	93.4

2.3.2 散射系数自动提取结果

自动提取散射系数结果如表 2 所示, 对比表 3 中三种方法的实验结果可以发现, 在模型精度方面, GRU-WH 的平均绝对误差和平均绝对百分比误差最高, 分别为 2.361 和 0.876; Unet-WH 的平均绝对误差和平均绝对百分比误差低于 GRU-WH, 分别为 1.532 和 0.324; GRU-MHA 的平均绝对误差和平均绝对百分比误差最低, 分别为 0.453 和 0.176. GRU-MHA 相比 Unet-WH 平均绝对误差和平均绝对百分比误差分别降低了 70% 和 46%. GRU-MHA 相比 GRU-WH 平均绝对误差和平均绝对百分比误差分别降低了 92% 和 80%. Unet-WH 方法和 GRU-WH 方法提取效果差的原因是不同的.

Unet-WH 方法提取效果差原因分析: Unet-WH 方法鲁棒性差, 出现将多个闪电哨声波事件识别为一个闪电哨声波事件或将一个哨声波事件识别为多个闪电哨声波事件的情况, 导致提取效果差. 使用

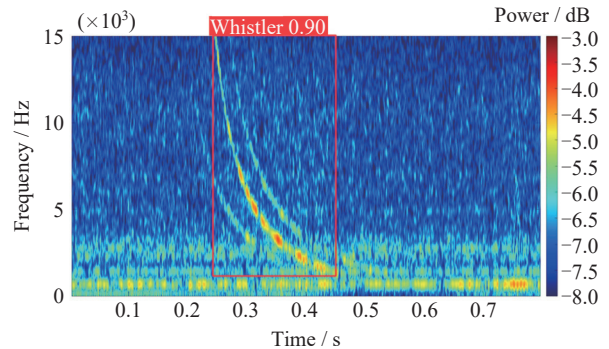


图 11 闪电哨声波漏检情况

Fig. 11 Lightning whistle leakage detection situation

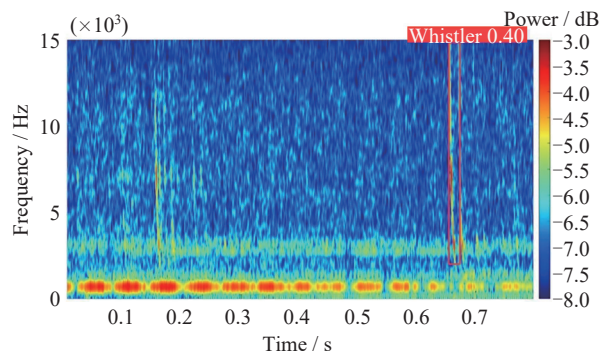


图 12 闪电哨声波误识别情况

Fig. 12 Diagram of misidentification of lightning whistle

GRU-MHA 方法和 Unet-WH 方法对分散和密集的闪电哨声波提取散射系数并与人工拟合结果对比, 结果如图 13 所示. Unet-WH 方法的鲁棒性差, 同时提取分散和密集的闪电哨声波时, 会出现较大偏差. GRU-MHA 方法的鲁棒性较好, 同时提取分散和密集的闪电哨声波时, 与人工拟合结果相差较小. GRU-WH 方法提取效果差的原因为闪电哨声波持续时间短, GRU 神经网络对哨声波散射系数提取能力不足, 导致提取效果差.

表 2 散射系数提取结果

Table 2 Scattering coefficient extraction results

性能指标	MAE	MAPE
人工方法	—	—
GRU-MHA	0.453	0.176
GRU-WH	2.361	0.876
Unet-WH	1.532	0.324

表 3 散射系数提取时间情况

Table 3 Scattering coefficient extraction time

性能指标	AET/s
人工方法	360
LWSC-AEM	0.074
Unet-WH	0.9

自动散射系数提取消耗时间情况如表 3 所示, 对比表 3 中两种方法的实验结果发现在时间消耗方面 Unet-WH 消耗时间长, 处理一张时频图的平均时间为 0.9 s; LWSC-AEM 处理一张时频图的平均时间为 0.074 s, 比 Unet-WH 减少了 92%. 原因是 Unet-WH 方法中使用的 Unet 神经网络计算成本高, 导致运行速度慢; 而 LWSC-AEM 方法中, YOLOV5 神经网络处理一张时频图的平均时间为 0.07 s, GRU-MHA 神经网络处理一张时频图的平均时间为 0.004 s.

通过以上结果分析, 本文提出的 LWSC-AEM 方法在精确度和时间消耗方面均优于 Unet-WH 方法.

2.4 模型定性评价

2.4.1 闪电哨声波识别结果

YOLOV5 识别闪电哨声波结果如图 14 所示. 观察图 14(a) 可以发现, 该闪电哨声波识别模型能够准确识别分散和密集的闪电哨声波; 同时, 观察图 14(b) 可以发现, 该模型也能够对连通性强和连通性弱的闪电哨声波进行准确识别; 观察图 14(c) 可以发现该识别模型对持续时间长或者持续时间短、频率跨度大或者频率跨度小的闪电哨声波都有很好的识别效果.

2.4.2 散射系数提取结果

图 15 给出了采用不同算法提取闪电哨声波散射系数的拟合结果. 紫线代表人工提取方法, 黑线代表

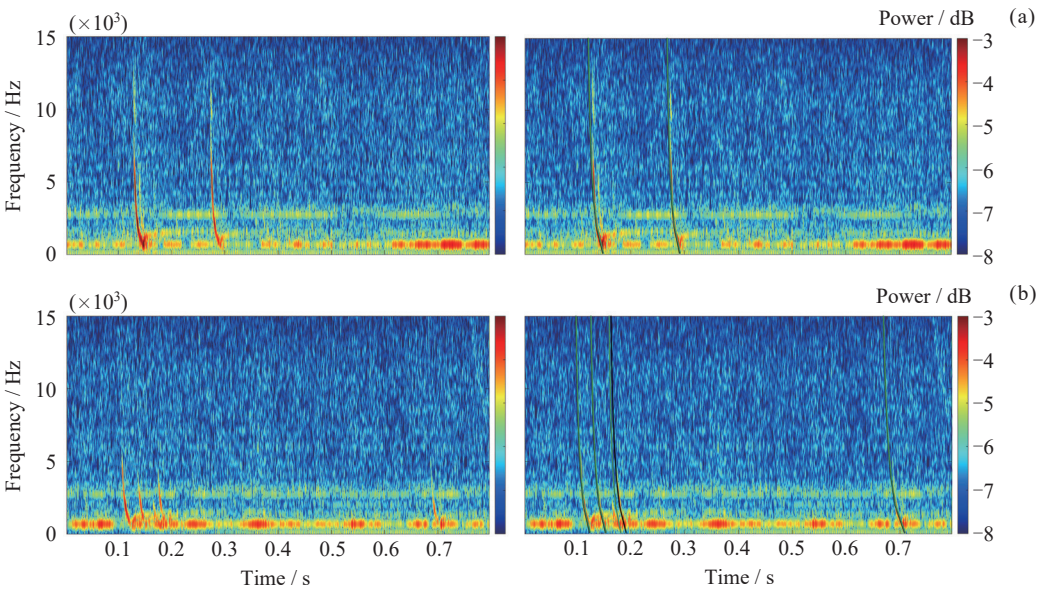


图 13 识别结果对比. (a) 分散的闪电哨声波, (b) 密集的闪电哨声波

Fig. 13 Comparison chart of recognition results. (a) Scattered lightning whistler, (b) dense lightning whistler

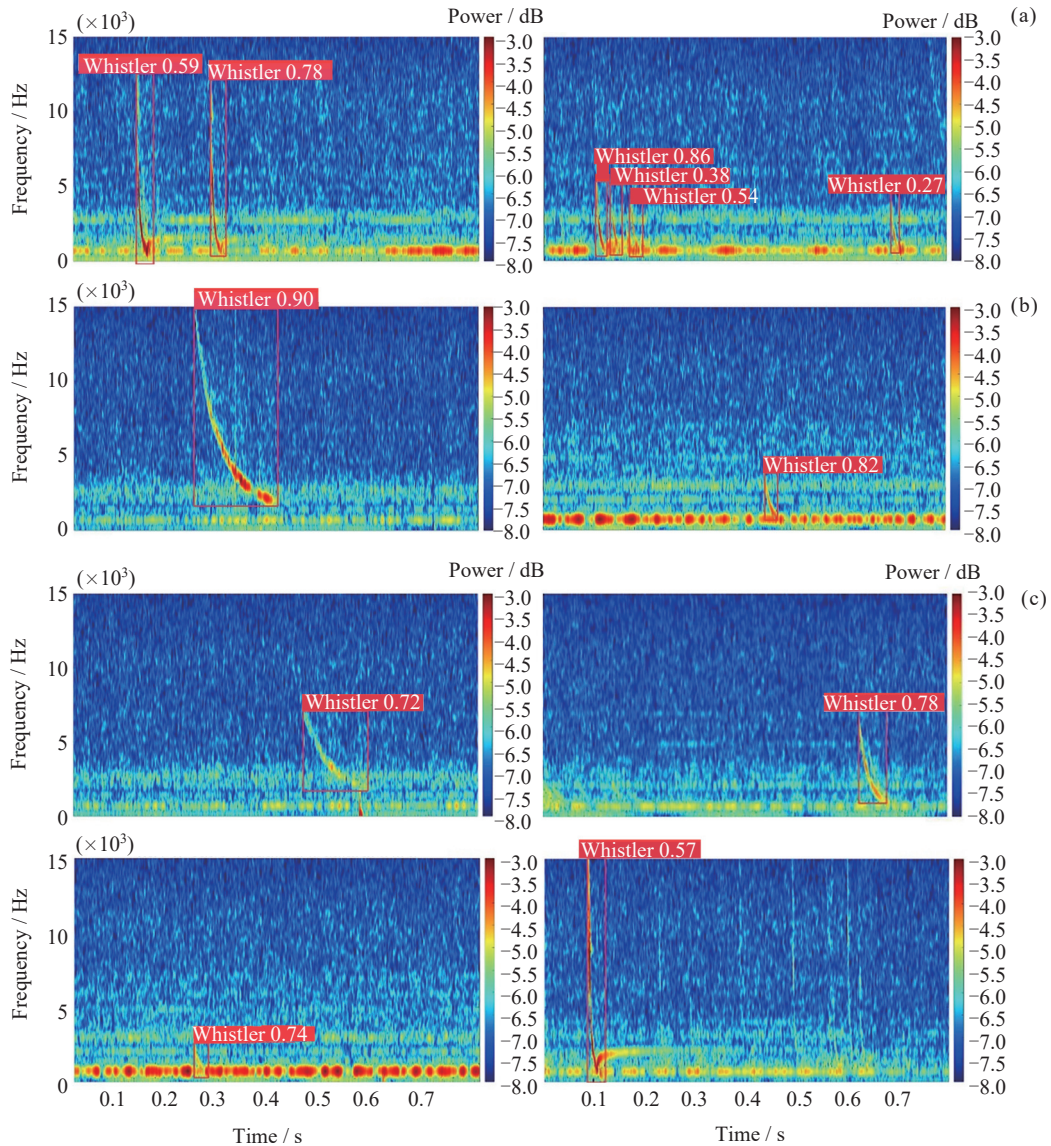


图 14 闪电哨声波识别结果. (a) 分散和密集的闪电哨声波, (b) 连通性强和连通性弱的闪电哨声波, (c) 持续时间长短或频率跨度不同的闪电哨声波

Fig. 14 Lightning whistler recognition rendering. (a) Scattered and dense lightning whistler, (b) lightning whistler with strong and weak connectivity, (c) lightning whistler with different durations or frequency spans

GRU-MHA 模型, 青线代表 GRU-WH 模型, 黄线代表 Unet-WH 模型. 观察图 15 发现, GRU-MHA 模型得到的闪电哨声波散射系数的拟合轨迹与人工方法相比差异较小, 表明该方法提取的散射系数更接近人工提取的散射系数; 而 GRU-WH 模型和 Unet-WH 模型的拟合效果较差.

图 16 为采用 GRU-MHA 方法提取不同散射系数的闪电哨声波的拟合结果. 紫线代表人工提取方法, 黑线代表 GRU-MHA 模型. 观察图 16 发现,

GRU-MHA 模型对不同散射系数的闪电哨声波提取结果与人工提取结果相差较小.

3 讨论

本算法的关键环节是基于多头注意力机制改进的 GRU 神经网络, 其优化器和多头注意力机制中的“头数量”对算法效果具有一定的影响.

3.1 优化器对结果的影响

评估 GRU-MHA 模型中五种不同优化器 (Adam,

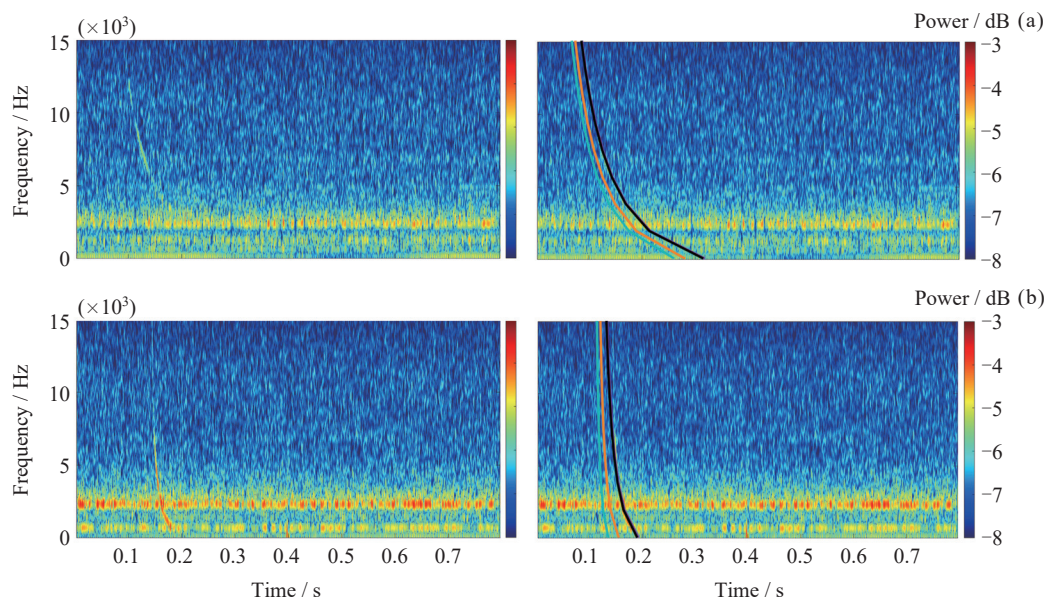


图 15 拟合效果
Fig. 15 Fitting effect

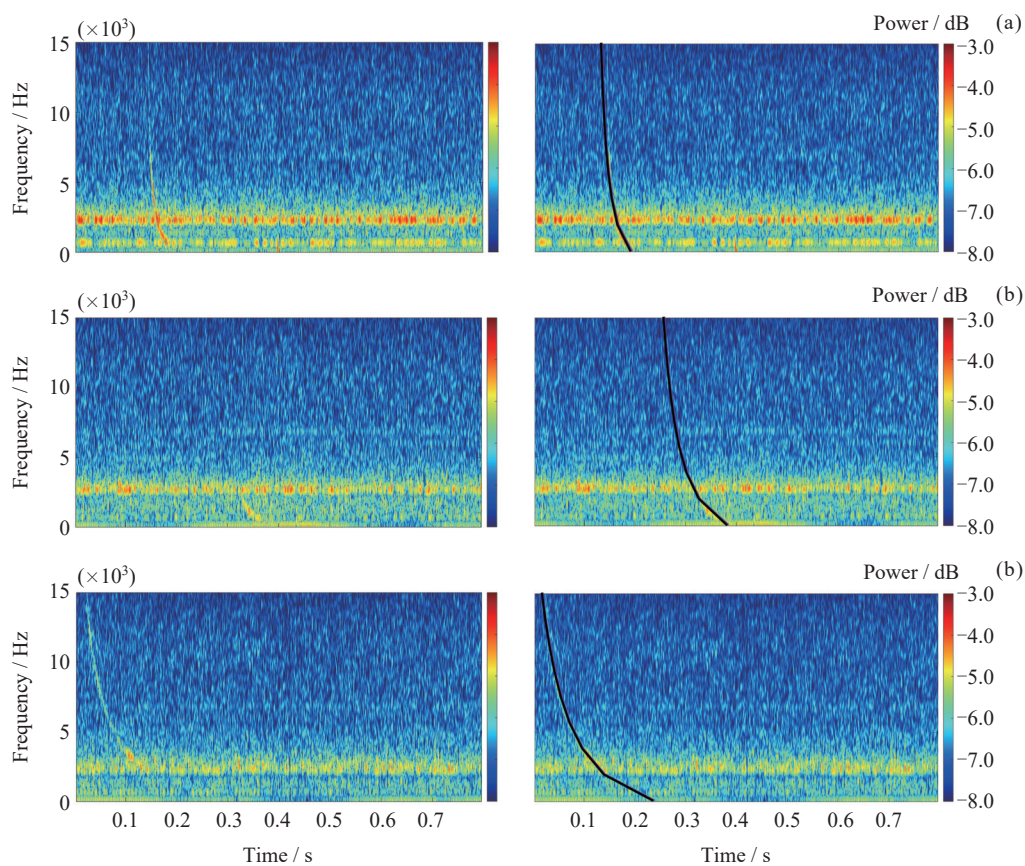


图 16 不同散射系数闪电电哨声波拟合结果
Fig. 16 Fitting results for lightning whistler with different scattering coefficients

Adamax, SGD, ASGD, PMSprop) 对算法效果的影响, 并绘制了折线图 (见图 17). Adam 和 Adamax 两个优化器呈现相似的效果和收敛趋势; 同样, SGD 和 ASGD 两个优化器的性能也相近. 然而, 在实验中 Adam 优化器效果优于其他优化器.

3.2 多头注意力机制中“头数量”对算法效果的影响

在 GRU-MHA 模型中, 多头注意力机制 (Multi-head Attention) 中“头数量”的选择对性能具有重要影响, “头数量”过大或过小都可能导致性能下降. 重点讨论“头数量”对 GRU-MHA 模型性能的影响. 因为需要将输入的特征均匀分配给每个“头”, 所以多头注意力机制的“头数量”必须能够整除输入特征数. 在本文中, 输入特征个数为 96, “头数量”的选择范围是 [1, 2, 4, 8, 16, 24, 32]. 进行了模型训练和性能评估, 实验结果如图 18 所示.

观察图 18 可以发现, 随着“头数量”的增加, 模

型大小也相应增加. 然而, 提取散射系数的准确度并未随着“头数量”的增加而持续提升. 分析其原因是多头注意力机制会将输入的特征 (96 个) 均匀分配给每个头, 增多的头数意味着每个头获得的特征数量减少, 导致每个头从每个特征中学到的信息增多. 然而, 过多的头数可能导致模型学习到不属于闪电哨声波的信息, 可能导致过拟合, 降低了散射系数提取效果. 比较分析结果显示, “头数量”为 16 时, 模型的提取效果最优且模型大小适中. 同时, “头数量”为 1 时, 模型提取效果与“头数量”为 16 时相近, 且模型大小最小. 此外, “头数量”为 1 时, 运行速度也更快.

不同“头数量”的闪电哨声波拟合效果如图 19 所示, 图 19(a) 为拟合整体图, 图 19(b) 为图 19(a) 黑色框中的局部图. 其中, 黑线代表人工拟合结果, 而不同“头数量”下的曲线分别为棕线 (头数量为 1), 黄线 (头数量为 2), 绿线 (头数量为 4), 紫线 (头数量为 8),

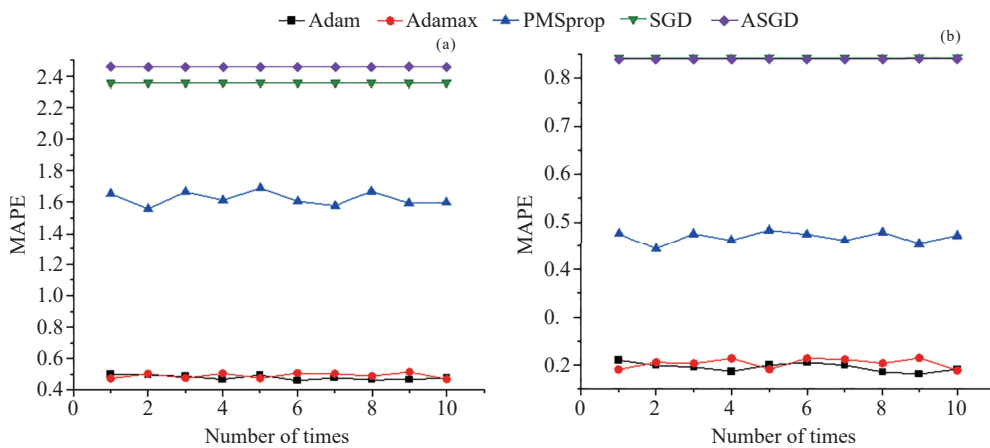


图 17 不同优化器 MAPE 和 MAE 效果对比

Fig. 17 Comparison of effects of different optimizers MAPE and MAE

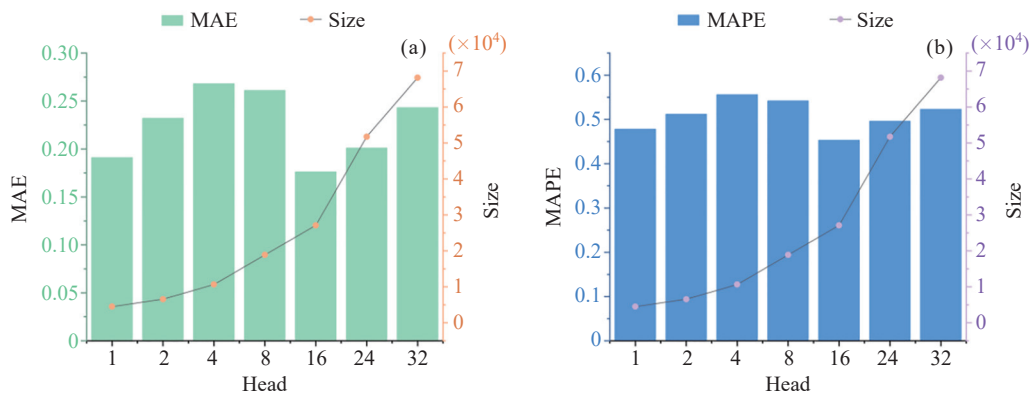


图 18 MAE 和 MAPE 与模型大小对比

Fig. 18 Comparison of MAE and MAPE with model size

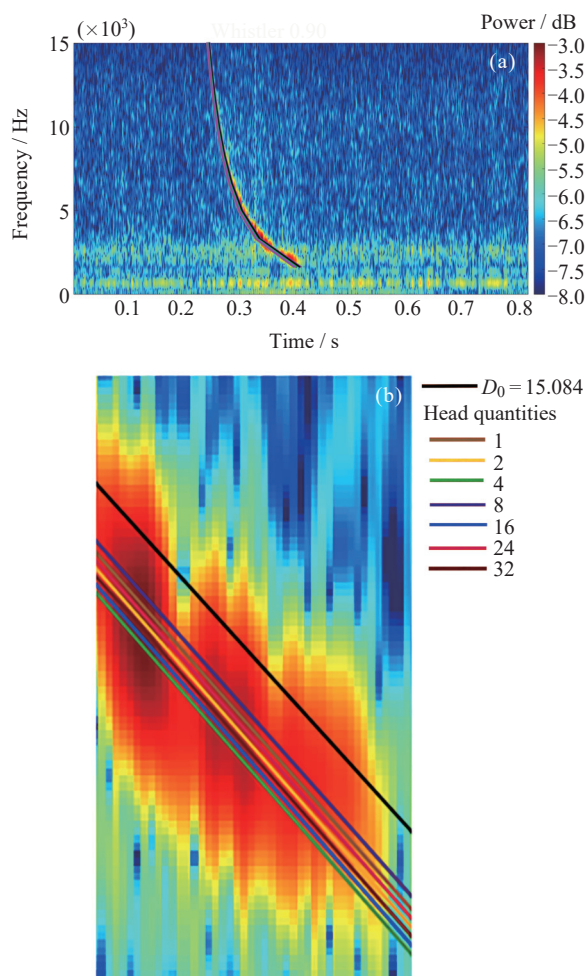


图 19 不同“头数量”的拟合效果

Fig. 19 Fitting effect of different head quantities

蓝线 (头数量为 16), 粉线 (头数量为 24), 以及深红线 (头数量为 32). 观察图 19(b) 发现, “头数量”为 16 的紫色拟合线最接近人工拟合的黑色拟合线. 不同头数量对于散射系数的提取效果相差很小, 这突显了 GRU-MHA 方法提取散射系数良好的稳定性.

4 结论

提出了一种快速稳健的闪电哨声波散射系数自动提取方法, 旨在解决 Unet-WH 方法鲁棒性差的问题. 该方法主要包括两个核心部分: 利用 YOLOV5 神经网络在时频图中识别闪电哨声波和采用多头注意力机制改进的 GRU 神经网络自动提取闪电哨声波音频散射系数. 基于张衡一号卫星 2020 年 2 月 SCM 载荷 VLF 波段的数据, 对该方法进行测试, 结果如下.

(1)YOLOV5 神经网络识别闪电哨声波的准确率

为 91.8%, 召回率为 93.4%, 漏检数少且精准度高, 保证了 GRU-MHA 方法提取散射系数的有效性.

(2)GRU-MHA 方法相较于 Unet-WH 方法在平均绝对误差和平均绝对百分比误差方面分别降低了 70% 和 46%.

(3)LWSC-AEM 方法相较于 Unet-WH 方法在使用时间方面减低了 92%.

综上所述, 本文提出的自动提取闪电哨声波方法相比于 Unet-WH 方法在速度和准确性上有较大的优势. 未来将继续提高 YOLOV5 的召回率和精确率, 降低 GRU-MHA 的平均绝对误差和平均绝对百分比误差, 完善智能音频自动提取闪电哨声波散射系数的方法.

致谢 研究数据来自国家自然灾害防治研究院的张衡一号卫星团队.

参考文献

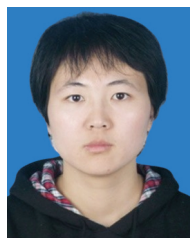
- [1] FENG Xiaokang, YUAN Jing, WANG Qiao, *et al.* Automatic extraction of physical parameters of lightning whistler recorded by search coil magnetometer onboard ZhangHeng Satellite[J/OL]. *Progress in Geophysics*, 2023, 38(6): 2373-2391
- [2] CHEN Y P, NI B B, GU X D, *et al.* First observations of low latitude whistlers using WHU ELF/VLF receiver system[J]. *Science China Technological Sciences*, 2017, 60(1): 166-174
- [3] CARPENTER D L, ANDERSON R R. An ISEE/whistler model of equatorial electron density in the magnetosphere[J]. *Journal of Geophysical Research: Space Physics*, 1992, 97(A2): 1097-1108
- [4] SINGH A K, VERMA U P, BHARGAWA A. Remote sensing of Mid/Upper atmosphere using ELF/VLF waves[J]. *Global Journal of Science Frontier Research: A Physics and Space Science*, 2018, 18(10): 10-21
- [5] 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*, 2014, 49(9): 753-764
- [6] 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
- [7] CLILVERD M A, NUNN D, LEV-TOV S J, *et al.* Determining the size of lightning-induced electron precipitation patches[J]. *Journal of Geophysical Research: Space Physics*, 2002, 107(A8): SIA 10-1-SIA 10-11
- [8] KISHORE A, DEO A, KUMAR S. Upper atmospheric remote sensing using ELF-VLF lightning generated tweek and whistler sferics[J]. *The South Pacific Journal of Natu-*

- ral and Applied sciences*, 2016, **34**(1): 12-20
- [9] PARROT M, PINÇON J L, SHKLYAR D. Short-fractional hop whistler rate observed by the low-altitude satellite DEMETER at the end of the solar cycle 23[J]. *Journal of Geophysical Research: Space Physics*, 2019, **124**(5): 3522-3531
- [10] HORNE R B, GLAUERT S A, MEREDITH N P, *et al.* Space weather impacts on satellites and forecasting the Earth's electron radiation belts with SPACECAST[J]. *Space Weather*, 2013, **11**(4): 169-186
- [11] REN Y, DAI L, LI W, *et al.* Whistler waves driven by field-aligned streaming electrons in the near-Earth magnetotail reconnection[J]. *Geophysical Research Letters*, 2019, **46**(10): 5045-5054
- [12] REN Y, DAI L, WANG C, *et al.* Statistical characteristics in the spectrum of whistler waves near the diffusion region of dayside magnetopause reconnection[J]. *Geophysical Research Letters*, 2021, **48**(1): e2020GL090816
- [13] ZHOU Huaibei, XIAO Zuo, SUN Chuanli. A study of cloud-earth lightning parameters by using whistler spectrum[J]. *Chinese Journal of Radio Science*, 1989(1): 1-7
- [14] HE Youwen. An analysis of whistler sources[J]. *Chinese Journal of Space Science*, 1981, **1**(2): 143-152 (何友文. 哨声“源”的分析 [J]. *空间科学学报*, 1981, **1**(2): 143-152)
- [15] ARTEMYEV A V, LAPSHIN N V, SIMONOV S A. Prospects of predicting the number of geese on spring migration stopover sites in Karelia[J]. *The Herald of Game Management*, 2014, **11**(2): 249-255
- [16] HU Yunpeng, ZHIMA Zeren, 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
- [17] HELLIWELL R A. Whistlers and Related Ionospheric Phenomena[M]. Stanford: Stanford University Press, 1965
- [18] LICHTENBERGER J, FERENCZ C, BODNÁR L, *et al.* Automatic whistler detector and analyzer system: automatic whistler detector[J]. *Journal of Geophysical Research: Space Physics*, 2008, **113**(A12): A12201
- [19] Stanford VLF Group. Automated detection of whistlers for the TARANIS spacecraft overview of the project[EB/OL]. 2009
- [20] DHARMA K S, BAYUPATI I P A, BUANA P W. Automatic lightning whistler detection using connected component labeling method[J]. *Journal of Theoretical and Applied Information Technology*, 2014, **66**(2): 638-645
- [21] KONAN O J E Y, MISHRA A K, LOTZ S. Machine learning techniques to detect and characterise whistler radio waves[OL]. arXiv preprint arXiv: 2002.01244v1, 2020
- [22] 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. DOI: 10.6038/cjg202100263
- [23] 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. DOI: 10.6038/cjg202100164
- [24] YUAN Jing, WANG Zijie, ZHIMA Zeren, *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. DOI: 10.6038/cjg2022P0365
- [25] ECKERSLEY T L. Musical atmospherics[J]. *Nature*, 1935, **135**(3403): 104-105
- [26] DAI Jian, ZHAO Xu, LI Lianpeng, *et al.* Improved YOLOV5-based infrared dim-small target detection under complex background[J]. *Infrared Technology*, 2022, **44**(5): 504-512
- [27] HE Yu, TIAN Junwei, ZHANG Zhen, *et al.* Lightweight research of YOLOV5 target detection[J]. *Computer Engineering and Applications*, 2023, **59**(1): 92-99
- [28] SHAO Yanhua, ZHANG Duo, CHU Hongyu, *et al.* A review of YOLO object detection based on deep learning[J]. *Journal of Electronics & Information Technology*, 2022, **44**(10): 3697-3708. DOI: 10.11999/JEIT210790
- [29] PANG Cong, JIANG Yong, LIAO Chengwang, *et al.* Automatic recognition of natural earthquakes and artificial blasting based on the sample entropy of the Mel frequency cepstrum coefficient and support vector machine optimized by gray wolf optimization[J]. *China Earthquake Engineering Journal*, 2022, **44**(5): 1169-1175
- [30] YUAN Zhengwu, XIAO Wanghui. Improved speech recognition algorithm based on MFCC feature[J]. *Computer Engineering and Applications*, 2009, **45**(33): 108-110 (袁正午, 肖旺辉. 改进的混合 MFCC 语音识别算法研究 [J]. *计算机工程与应用*, 2009, **45**(33): 108-110)
- [31] TAN Xiangyong, HU Tianying, LIU Feng. Serial correlation test for linear models under Huber loss[J]. *Journal of Chongqing University of Technology (Natural Science)*, 2023, **37**(8): 342-347

作者简介



韩金昇 男, 1999 年 1 月出生于河北省唐山市, 硕士。主要研究方向为地球观测数据智能处理与分析等领域。
E-mail: 18812162636@163.com



袁静 (通信作者) 女, 1981 年 3 月出生于河北省石家庄市, 现为防灾科技学院副教授, 硕士生导师, 主要研究方向为地球观测数据智能处理与分析等领域。
E-mail: yuanjing20110824@sina.com