

WANG Heye, GUO Xuedou, ZHANG Sai, HUANG Yangyang, LIU Tianle, ZHAO Shuya. Recognition Method of Auroral Kilometric Radiation Based on Fusion of VGG16 and Self-attention Mechanism (in Chinese). *Chinese Journal of Space Science*, 2025, 45(3): 677–688. CSTR:32142.14.cjss.2024-0043. DOI:10.11728/cjss2025.03.2024-0043

基于 VGG16 与自注意力机制融合的 极光千米波识别

王鹤野¹ 郭雪豆¹ 张赛¹
黄泱泱¹ 刘天乐² 赵舒雅¹

1(长沙理工大学物理与电子科学学院 长沙 410114)

2(长沙理工大学电气与信息工程学院 长沙 410114)

摘要 提出了一种能准确识别极光千米波 (Auroral Kilometric Radiation, AKR) 的方法, 为进一步研究极光千米波在地球辐射带能量粒子剧烈变化过程中的作用提供支撑. 首先采用 VGG16 卷积神经网络作为基础模型, 从原始数据中提取出有助于识别 AKR 的局部特征. 之后引入嵌入 VGG16 网络的定制化自注意力机制 (Self-Attention Mechanism embedded in VGG network, SAM-V), 该机制有助于捕捉功率谱图中不同时间点或频率上的信号可能存在的关联, 减小其他杂波的影响, 提高识别准确性. 同时, 采用线性学习率预热和动态衰减策略使模型更快地收敛并提高泛化能力. 实验结果表明, 改进后的模型平均识别准确率在 93% 左右, 比原始 VGG16 平均提高约 3.3%, 并且召回率和精确度等指标均有所改善.

关键词 VGG16 卷积神经网络, 自注意力机制, 极光千米波, 学习率策略

中图分类号 TP751

Recognition Method of Auroral Kilometric Radiation Based on Fusion of VGG16 and Self-attention Mechanism

WANG Heye¹ GUO Xuedou¹ ZHANG Sai¹
HUANG Yangyang¹ LIU Tianle² ZHAO Shuya¹

1(School of Physics & Electronic Science, Changsha University of Science & Technology, Changsha 410114)

2(School of Electrical & Information Engineering, Changsha University of Science & Technology, Changsha 410114)

Abstract An enhanced image recognition algorithm for Auroral Kilometric Radiation (AKR) detection is presented by integrating a self-attention mechanism into the VGG16 Convolutional Neural Network (CNN) architecture. The primary goal is to improve the flexibility and detection accuracy of AKR

2024-03-17 收到原稿, 2024-07-31 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/>)

identification, which is crucial for understanding the dynamic changes in Earth's radiation belts and the associated energetic particle variations. The methodology begins with employing the VGG16 CNN as the foundational model to extract local features from raw data that are instrumental in AKR recognition. Subsequently, a custom Self-Attention Mechanism (SAM-V) is embedded in the VGG network. The Self-Attention Mechanism (SAM), originally designed for sequential data processing, is adapted to work with the VGG16 network. Traditional integration of SAM with VGG16 could potentially increase the model's complexity and computational cost, leading to potential feature sparsity issues. However, the proposed custom SAM-V generates queries, keys, and values through defined convolutional layers, offering more control over feature input and output. This customization implies shared parameters, reducing the number of model parameters, thereby mitigating the risk of overfitting and enhancing the model's generalization capabilities. This approach is particularly adept at capturing correlations in power spectral density across different time points or frequencies, minimizing the impact of noise and improving recognition accuracy. Additionally, a linear learning rate warm-up and dynamic decay strategy are employed to accelerate model convergence and enhance generalization. The experimental results demonstrate that the improved model achieves an average recognition accuracy of approximately 93%, which represents a 3.3% increase compared to the original VGG16 model. Furthermore, other performance metrics such as recall rate and precision have also seen significant improvements. In conclusion, the integration of a custom self-attention mechanism into the VGG16 network has yielded a more efficient and accurate model for AKR detection. This advancement not only bolsters the study of auroral kilometric radiation but also has broader implications for the analysis of Earth's radiation belt dynamics and energetic particle behavior. The model's enhanced generalization capabilities and improved accuracy underscore the potential for applying similar techniques to other image recognition tasks within the field of space physics and beyond.

Key words VGG16 convolutional neural network, Self-attention mechanism, Auroral kilometric radiation, Learning rate policy

0 引言

极光千米波是一种在地球磁层中广泛存在的高频电磁波,与地球的分立极光存在密切的关联^[1]. 1974年 Gurnett 等^[2]研究了极光千米波与地球极光的关系,并将其命名为 TKR (Terrestrial Kilometric Radiation); 1975年 Kurth 等^[3]将其称为 AKR (Auroral Kilometric Radiation). AKR 的主要特征^[4,5]包括: 频率范围 30~800 kHz; 功率可高达 1000000 kW; 波长约为 1000 m; 模式主要是右旋异常模,少数为左旋寻常模和左旋异常模. 理论研究表明^[6], 极光千米波也是一种能够与辐射带电子产生回旋共振的等离子体波动. 这种波粒相互作用能够引起辐射带高能电子通量急剧增加或减少,从而对在轨航天器造成严重危害或者影响大气成分变化^[7], 因此对极光千米波的识别在具有重要研究意义.

VGG 是一种深度卷积神经网络,由牛津大学视觉几何组 (Visual Geometry Group) 于 2014 年提

出^[8]. VGG 是由多个卷积层和池化层组成的神经网络,具有很强的图像分类能力. 2018 年 Wang 等^[9]采用深度学习的方法,利用卷积神经网络模型自主表征极光特征,将其应用于极光图像分类并取得较高的准确率. 2019 年 Li 等^[10]提出一种改进的 Dense_VGG 网络,实现了极光图像的自动分类,并且提高了准确率. 2020 年 Niu 等^[11]提出了基于深度学习的极光局部结构识别与定位. 2022 年 Ren 等^[12]基于深度学习的方法,提出了极光事件的自动识别与检测. 深度学习正逐渐成为研究空天观测相关问题的重要手段.

本文根据极光千米波的特征,以 VGG 卷积神经网络为基础,提出一种灵活性更强、检测精度更高的 VGG16 融合自注意力机制的图像识别算法. 自注意力机制 (Self-Attention Mechanism, SAM)^[13]是一种用于处理序列数据的机制,传统的自注意力机制与 VGG16 融合可能会增加模型的复杂度和计算成本,进而导致特征稀疏性问题. 嵌入 VGG-net 的定制化自注意力机制 (SAM-V) 可通过自定义卷积层来生成

查询、键和值, 更灵活地控制特征输入和输出, 这意味着部分参数是共享的, 因而可以减少模型参数的数量, 从而降低过拟合的风险, 并提高模型的泛化能力.

本文主要基于 VGG16 网络提取图像的局部特征, 使用 SAM-V 关注图像不同时域中更为关键和显著的频域和功率谱密度特征, 使模型更好地捕捉 AKR 观测图像中的长距离依赖关系^[14], 从而提高模型的性能. 在训练过程中, 采用基于 AKR 特征的学习率预热^[15]和步进衰减策略, 同时自注意力机制会与其他神经网络层一起进行优化, 以学习到更为有效的权重分配策略.

1 基于 VGG16 网络的 AKR 识别

1.1 网络架构

VGG16 网络由 13 个卷积层 (Convolutional Layers)、3 个全连接层 (Fully Connected Layers) 和 5 个最大池化层 (Max Pooling Layers) 组成. 网络输

入为 224×224 Pixel 的彩色图像, 经过一系列卷积和池化操作后, 特征图的尺寸逐渐减小而深度逐渐增加. 每个卷积层后面都紧跟一个 ReLU 激活函数, 用于引入非线性特性. 网络深度和宽度的逐渐增加会使得网络能够捕捉到不同尺度和层次的特征. 在进入全连接层之前, 特征图被展平为一维向量, 然后传递给全连接层进行分类. VGG16 网络的执行过程如图 1 所示, 部分细节如图 2 所示.

1.2 细节处理

在卷积层中, 卷积核连接的权重向量 (Weight) 和偏置向量 (Bias) 会在运算时加到卷积结果上, 其输出按行堆叠形成更大的特征图. 采取平均池化操作, 即保留窗口中的平均值. 在多个池化核并行工作的情况下可以提取输入图像的多个维度信息, 同时 Avg-pooling 能够保留图像的纹理信息并减小卷积层参数误差造成的估计偏移. ReLU (修正线性单元) 的主要作用是神经网络引入非线性, 增强模型的非线性表达能力. ReLU 对负值直接截断, 避免了梯度爆炸问

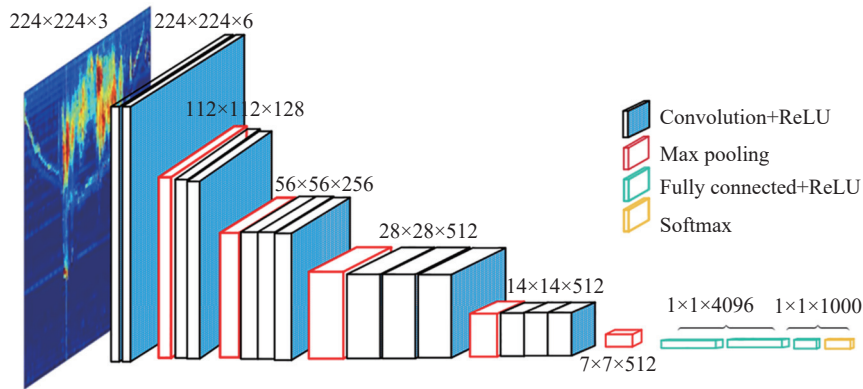


图 1 VGG16 网络结构

Fig. 1 Network structure of VGG16

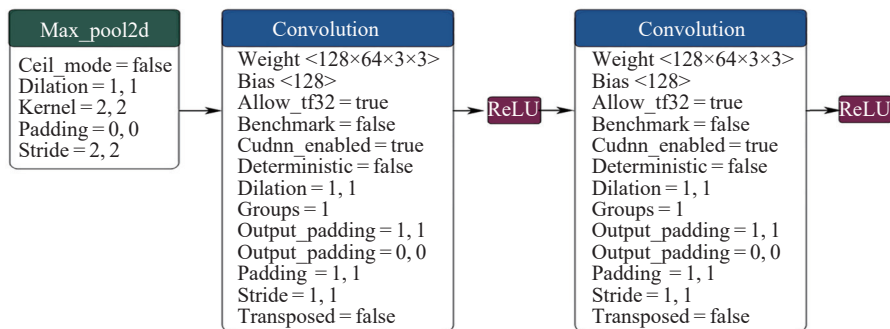


图 2 VGG16 执行过程部分细节

Fig. 2 Partial details of the VGG16 execution process

题. 图 3 给出了使用 3×3 卷积核、 2×2 池化核以及非线性激活后各个通道的形象.

全连接层 (Fully Connected, FC) 是网络的最后一部分, 其通过连接线上的权值进行前向传播, 加权组合得到下一层神经元节点的输入^[16,17]. 全连接层中的每一层是由许多神经元组成的平铺结构, 其后接一个 ReLU 激活函数和一个 Dropout 层以减少过拟合. 这些组合有助于模型从复杂的图像特征中学习抽象的概念, 并最终生成每个类别的得分. 图 4 给出了一个全连接层转换示意.

在 VGG16 网络识别 AKR 过程中, 输入为 $224\times 224\times 3$ 分割的电磁场功率谱图, 最后一层卷积可得输出为 $7\times 7\times 512$, 如果后一层是包含 4096 个神经元的 FC, 则其可用 $7\times 7\times 512\times 4096$ 的卷积核通过全局卷积来实现 FC 运算过程. 在 FC 的这一步卷积操作中, 其核心功能是将分散在不同空间位置的特征表示汇聚到统一的表征空间中, 从而将多维特征信息融合为单一的数值输出. 这种转换极大地降低了特征位置差异对分类结果的影响, 即不同的时域范围对于识别 AKR 本身无影响.

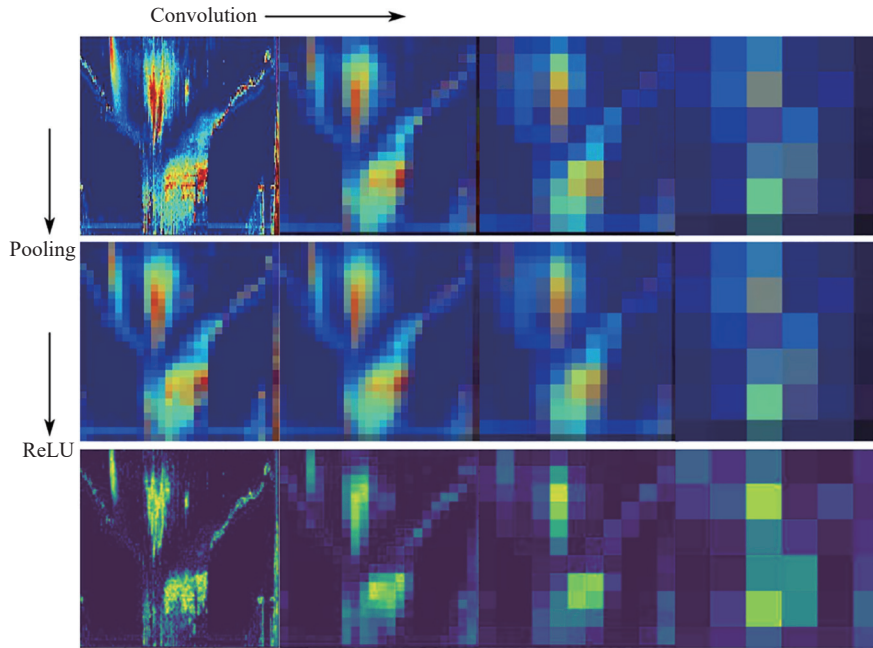


图 3 经过卷积、池化层、激活函数后的变化

Fig. 3 Changes after convolution, pooling layer and ReLU activation function

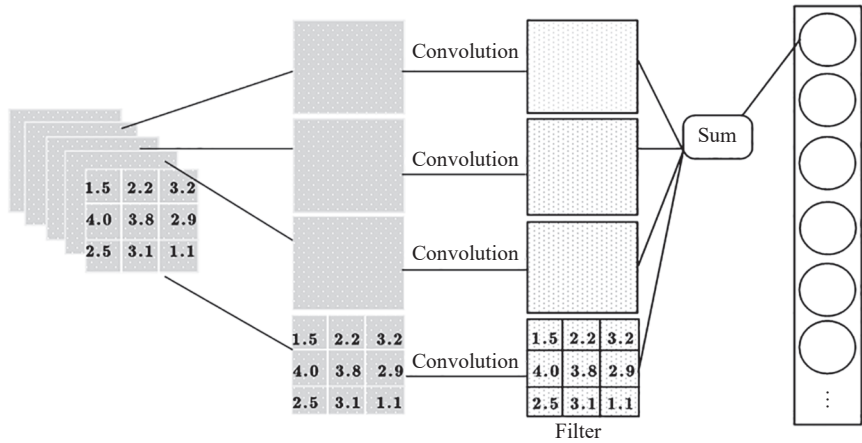


图 4 $3\times 3\times 4$ 经全连接层转换为 1×4096

Fig. 4 $3\times 3\times 4$ is converted to 1×4096 by the fully connected layer

1.3 优化器与损失函数

采用 Adam 优化器来更新网络参数, 其结合了 SGD 与 AdaDelta 算法的优点^[18], 通过计算梯度的一阶矩估计和二阶矩估计来调整每个参数的学习率^[19], 同时还包括了动量项和平方梯度项的偏差校正, 以提高收敛速度和稳定性. 在分类任务中, 由于 AKR 图像的多样性, 模型需要捕捉不同时刻的功率谱和空间特征. Adam 优化器通过为每个参数动态调整学习率^[20], 帮助模型快速收敛, 同时避免陷入局部最小值. 最常见的两个超参数 β_1, β_2 出现在下式, 前者为 SGD 的一阶动量, 后者为控制 AdaDelta 的二阶动量.

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) g_t, \quad (1)$$

$$\mathbf{V}_t = \beta_2 \times \mathbf{V}_{t-1} + (1 - \beta_2) g_t^2. \quad (2)$$

本文使用了交叉熵损失函数, 其可衡量实际标签的概率分布与模型预测的概率分布之间的差异^[21], 能够有效指导模型学习到区分 AKR 的重要时频特征, 同时帮助模型优化对 AKR 分类的决策边界, 从而提高对 AKR 识别的准确性. 在 PyTorch 中, CrossEntropyLoss 函数自动结合了 LogSoftmax 与负对数似然损失, 模型能够更好地泛化到不同磁纬度观察的 AKR 图像, 即使在不同观测条件和时域下也能够保持高识别率.

对 AKR 二分类得到的概率为 p 和 $1-p$, 此时表达式如下:

$$L = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i -[y_i \cdot \ln(p_i) + (1 - y_i) \ln(1 - p_i)]. \quad (3)$$

式中, y_i 表示样本 i 的 label, 正类为 1, 负类为 0; p_i 表示样本 i 预测为正类的概率.

2 融入自注意力机制

2.1 AKR 形态与特征

早期的极区 AKR 观测主要集中在极地地区. 范艾伦探测器的轨道在南北半球 20° 的纬度范围内, Kurth 等^[22]2015 年通过范艾伦探测器的观测数据揭示了在辐射带赤道区域也存在 AKR. 本文所使用的数据是 2012 年 10 月至 2019 年 7 月期间范艾伦探测器的部分观测数据.

图 5 为两组观测数据的电磁场功率谱密度, 其频率范围为 $10 \sim 400$ kHz. 其中频率范围较宽 ($50 \sim 800$ kHz), 峰值波频 f_p 变化较快, 频率大于上混杂 f_{cut} , $\text{PSD} > 10^{-14} \text{ V}^2 \cdot \text{m}^{-2} \cdot \text{Hz}^{-1}$ 且波长在千米级别的波称为极光千米波 (AKR). 在识别过程中要剔除频率分布过宽、功率谱连续过强的部分 (例如射电暴),

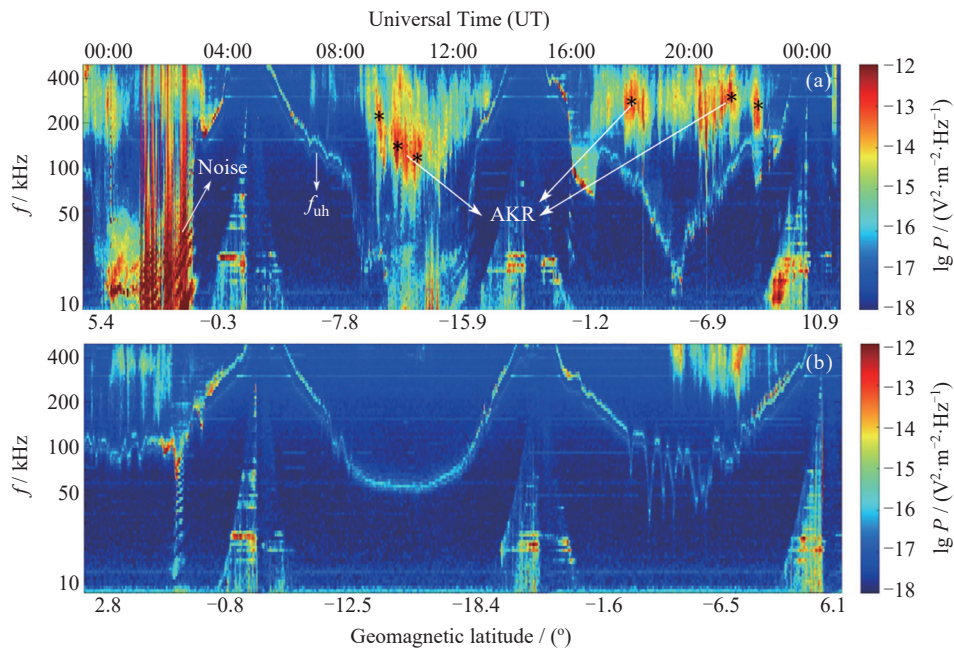


图 5 电磁场功率谱密度

Fig. 5 Power spectral density of electromagnetic field

以及频率处于 AKR 频率以下且 PSD 过小的部分 (高杂波、ECH 等), 这些被认为是来自极光区域的千米波辐射的背景噪声, 因此引入自注意力机制是十分必要的. 图 5(a) 与 (b) 相比, 可以观测到更多的 AKR, 分为 6 个频率中心. 经过融入定制化自注意力机制的 VGG16 模型处理后, 可判断不同时刻不同磁纬度是否存在 AKR; 随后可经过聚类与灰度匹配的图像处理来获取频率中心点的功率谱强度. 标记图 5(a) 中信息如表 1 所示.

这里更关注极光千米波的自动识别. AKR 向内磁层的传播会对磁层中粒子的演化产生部分影响. 因此, 准确识别出极光千米波是求解其时空分布特性的重要基础, 该方法也可用于验证新的行星数据中是否存在 AKR 的情况.

2.2 定制自注意力机制原理

在传统的卷积神经网络 (CNN) 中, 信息在序列中的传播是顺序的, 这可能导致长距离依赖问题, 即序列中远距离的元素之间的相互影响难以捕捉. 同时 AKR 观测具有时空上的复杂性, 信号可能在较长的时间、频率、功率谱尺度上具有相关性. 自注意力机

制 (SAM-V) 通过计算序列中每个元素与其他所有元素的关联强度来解决此问题. 具体来说, SAM-V 机制包括以下步骤 (见图 6).

(1) 计算 Quer, Key, Value: 通过对输入序列中每个元素矩阵变换得到其对应的 Query, Key 和 Value 三个向量.

(2) 计算注意力得分: 计算序列中的任意两个元素的 Query 与其他所有元素的 Key 的点积, 其结果表明了元素之间的关联强度.

(3) 应用 softmax 函数: 将上一步计算的得分进行 softmax 归一化后, 其可以被解释为每个元素对当前元素的贡献概率.

(4) 计算加权的 Value: 用 softmax 得到的概率对 Value 进行加权求和, 结果融合了序列中其他所有元素的贡献.

(5) 输出合并: 对已加权的 Value 进行线性变换, 得到 Self-Attention 的最终输出.

2.3 定制自注意力机制与 VGG16 网络的融合

定义 VGGWithAttention 与 Self-Attention 注意力模块. 前者接受一个特征图作为输入并与 VGG 网络连接, 后者输出一个加权的特征图; 后者模块包含 3 个卷积层: Query_conv, Key_conv 和 Value_conv. 这些卷积层分别用于提取查询、键和值的特征. 其被添加在特征提取网络 (Features) 和平均池化层 (Avgpool) 之后, 紧接在网络的分类层 (Classifier) 之前, 过程如图 7 所示.

提取 (Features): 使用 VGG16 对输入的图像进

表 1 AKR 频率中心点的部分信息

Table 1 Partial information about the AKR frequency center

	1*	2*	3*	4*	5*	6*
UT	09:30	10:00	11:10	18:20	21:30	22:00
$\lg P / (V^2 \cdot m^{-2} \cdot Hz^{-1})$	-12.55	-12.05	-12.75	-13.05	-12.65	-12.90

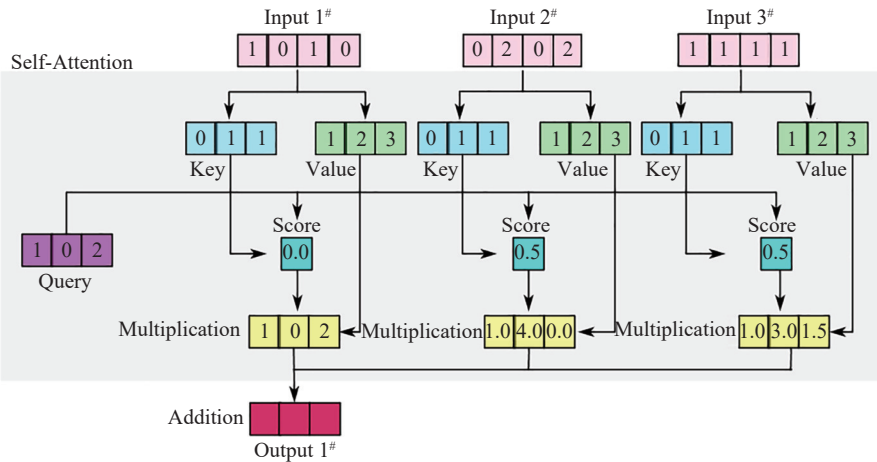


图 6 自注意力机制的计算过程
Fig. 6 Computation process of self-attention mechanism

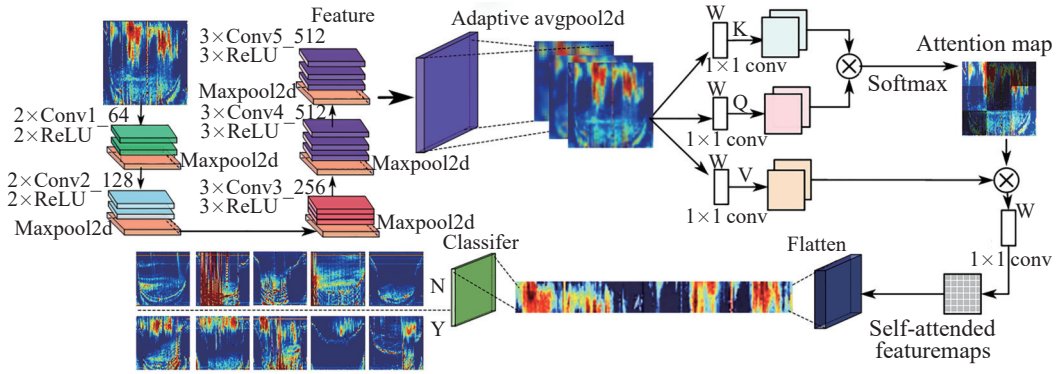


图 7 融合后的整体网络

Fig. 7 Overall network after fusion

行特征提取.

自注意力机制: 输出特征张量传递给 SAM-V. 机制经过自适应的处理, 对特征图中每个位置与其他所有位置的关系进行建模, 得到一个注意力权重矩阵.

加权特征融合: 利用注意力权重对特征图中每个位置的特征进行加权融合, 使每个位置的输出特征都包含图像中其他位置的上下文信息.

传播: 处理后的特征图通过 self.avgpool 与 torch.flatten 进行池化和展平操作.

分类层: 将经过自注意力加权融合后的特征输入到全连接层进行分类. 这些全连接层将特征映射到类别概率分布上^[23], 通过 softmax 函数得到每个类别的概率.

2.4 学习率预热与步进衰减

学习率预热的目的是让网络适应学习过程, 常见的预热方法包括线性预热、余弦预热、阶梯预热等. 这里采用线性预热帮助模型更快地学习到 AKR 观测图像的关键特征, 并避免在训练初期由于学习率过小导致的训练停滞. 定义如下关系式:

$$l_r = l_r + \frac{l_{r,\max} - l_{r,\min}}{N_{\text{epochs}} - N_{\text{warmup}}} \times (t - N_{\text{warmup}}). \quad (4)$$

其中, $l_{r,\min}$ 为最小学习率, $l_{r,\max}$ 为最大学习率, t 为当前轮数, N_{warmup} 为预热期长度.

步进衰减在训练后期通过减小学习率来减少模型对训练数据的过度依赖, 提高模型的泛化能力并平衡训练速度和模型性能. 常见的步进方法包括指数步进、阶梯步进、动态步进等. 本文采用动态步长衰减, 即学习率衰减的步长可以根据训练过程中 AKR 识别

的损失函数值进行动态调整. 定义函数关系如下:

$$l_r = l_r \times \left[1 - \frac{f(t)}{f(t-1)} \right]. \quad (5)$$

其中 $f(t)$ 为当前轮数的损失函数值.

3 结果与分析

3.1 数据预处理与实验环境

数据源于 2012 年 10 月至 2019 年 7 月期间范艾伦探测器工作的部分数据. 根据 Zhang 等^[24-26] 前期的工作, 通过程序自动筛选和人工分类整理了某段时间里事件库的数据集, 共计 24852 张, 其中标签 N 共 18642 张, 标签 Y 共 6210 张. 采用此数据集验证 VGG16 与 SAM-V 融合后产生的识别优化效果, 使用 DataLoader 类进行批量加载, 经过数据清洗后所有图像被转换为 PyTorch 张量格式并移动到 GPU 上. 为增强模型的泛化能力, 对图片应用随机裁剪、水平翻转、垂直翻转和随机旋转. 通过维度调整将所有图像调整为统一的尺寸 (图像数量 N , 图像通道数 C , 图像空间尺寸 H 和 W), 以匹配模型的输入层要求. 数据集被划分为训练集和测试集, 其中随机产生 10% 的数据 (2485 张) 用于测试. 实验环境配置见表 2.

3.2 评价系统

混淆矩阵: 表现形式如图 8 所示. 在图 8 中 TP, FP, FN, TN 分别表示被正确预测的正例、被错误预测的正例、被错误预测的反例、被正确预测的反例, 定义数学符号分别为 N_{TP} , N_{FP} , N_{FN} , N_{TN} .

准确率 (Accuracy, 定义符号 R_{accuracy})^[27]: 正确预测的样本数占总样本数的比例.

$$R_{\text{accuracy}} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}} + N_{\text{TN}}}.$$
 (6)

精确率 (Precision, 定义符号 $R_{\text{precision}}$): 即真正例占有所有预测为正类别的比例. 精确率越高, 说明模型对负样本的区分能力越强.

$$R_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}.$$
 (7)

召回率 (Recall, 定义符号 R_{recall})^[28]: 正类别样本被正确预测为正类别的比例. 召回率越高, 说明模型对正样本的识别能力越强.

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}.$$
 (8)

表 2 实验环境配置

Table 2 Configuration of the experimental environment

GPU	CPU	架构	环境
RTX 3090	AMD EPYC	PyTorch 1.7.0	Python3.8 Cuda11.0

	Actual positive	Actual negative
Predicted positive	TP (True Positive)	FP (False Positive)
Predicted negative	FN (False Negative)	TN (True Negative)

图 8 混淆矩阵

Fig. 8 Diagram of the confusion matrix

F1 分数 (F1 score, 定义符号 F_1): 精确率和召回率的调和平均数, F1 score 越高, 说明模型越稳健.

$$F_1 = \frac{2 N_{\text{precision}} N_{\text{recall}}}{N_{\text{precision}} + N_{\text{recall}}}.$$
 (9)

ROC 曲线: 不同阈值下分类器的真正率 (TPR, 定义符号 R_{TP}) 和假正率 (FPR, 定义符号 R_{FP}) 之间的关系. AUC 值表示 ROC 曲线下的面积, 用于衡量分类器性能. AUC 值越接近 1, 表示分类器性能越好.

$$R_{\text{TP}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}},$$

$$R_{\text{FP}} = \frac{N_{\text{FP}}}{N_{\text{FP}} + N_{\text{TN}}}.$$
 (10)

指标与置信度评价曲线: 在低置信度阈值下, 模型可能会过于宽松, 导致召回率很高而精确度下降; 在高置信度阈值下, 模型可能会过于严格, 导致精确度很高而召回率下降.

3.3 对比与分析

原始网络是指仅使用 VGG16 和迁移权重的训练模型. 改善后网络是指在 VGG16 基础网络上使用预训练权重 (迁移学习), 并融入 SAM-V, 引入线性学习率预热和动态衰减策略. 表 3 是两种模型的参数, 最终两种方式训练结果如图 9 所示. 图 9 中给出了两种

表 3 模型参数的数据对比

Table 3 Comparison of model parameters and other data

NET	Parameters	FLOPs	Model_size/Byte
VGG16-net	134268738	15466226688.0	539190670
VGG16+SAM-V	134794051	15491916800.0	553441679

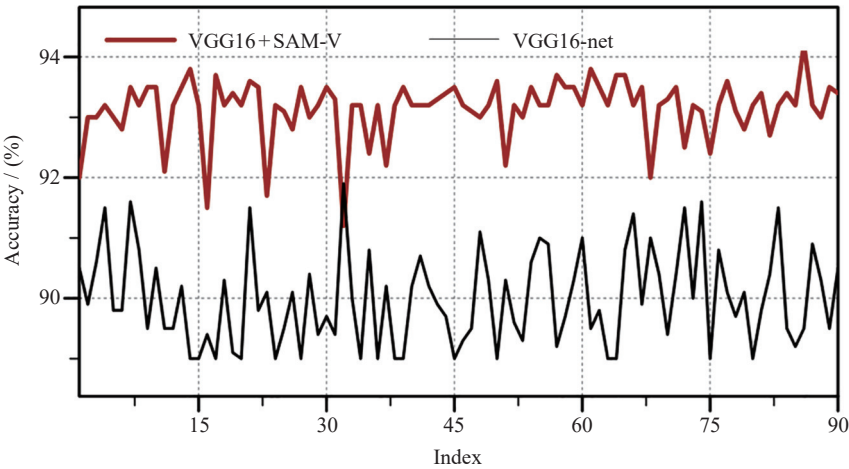


图 9 模型准确率折线对比

Fig. 9 Diagram of the confusion matrix

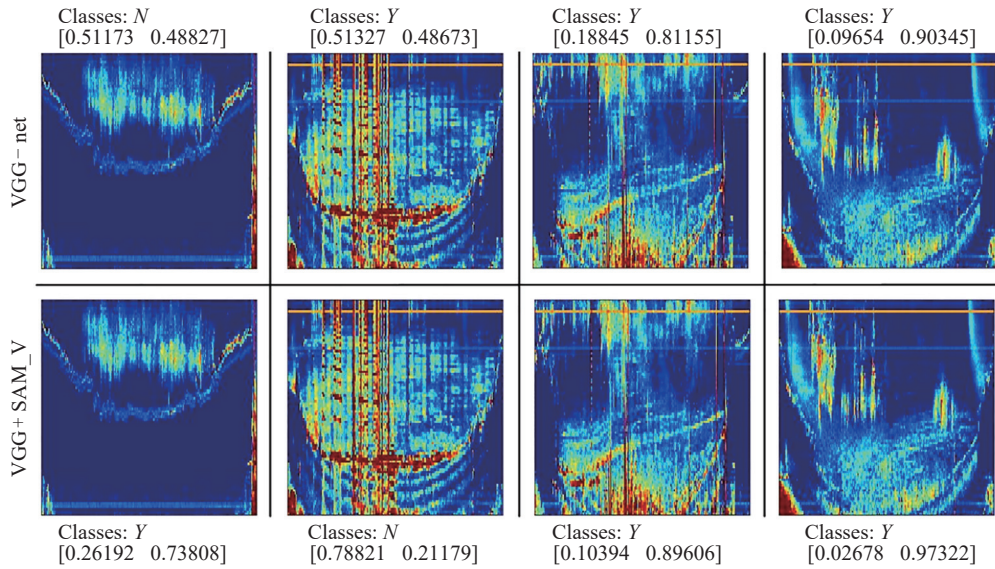


图 10 部分数据分类结果对比

Fig. 10 Comparison of some data classification results

模型预测的准确率随迭代次数的变化, 后者整体提高大于 3.3%.

图 10 是 AKR 特征较弱的数据识别结果, 相对于正常的分类测试, 上述图集因高杂波、ECH 波、射电暴等噪声影响而易被误识别. 相比原始 VGG16 的网络, 融入了定制化自注意力机制的模型识别类别均正确, 在此基础上分类概率值也较高.

图 11 给出了两种网络的置信度评价, 由图 11(a)(b) 可以看出, 相比原始模型, 改善后的网络中 Accuracy, Precision, Recall, F1 score 等随置信度的改变而变换的更平缓 and 紧促, 而且在置信度 0.2~0.8 内各评价价值均较高, 模型具有较好的泛化能力. 在置信度 0.7 左右时出现交汇点, 使得模型在精确度、召回率和准确率之间取得较好的平衡.

由图 12(a)(b) 可以观察到, 与原始网络训练结果相比, 改善后的网络 ROC 曲线、F1 曲线以及召回率曲线的 AUC 都更接近 1, 对于正类的预测有着很高的准确性, 并且能够很好地捕捉到正类样本; 在阈值较为宽泛的区间内, F1 得分、召回率以及精确率值均更高, 这表明模型能够在确保较少漏掉阳性样本的同时, 保持较高的预测准确性.

图 13 是模型训练过程中产生的混淆矩阵. 由图 13 可看出, 改善后网络的 TN 和 TP 占比均有所上升, FN 和 FP 占比均减少, 并且对于正类和负类的识别都有所提高, 模型的整体性能更佳, 分类更加准确和可靠.

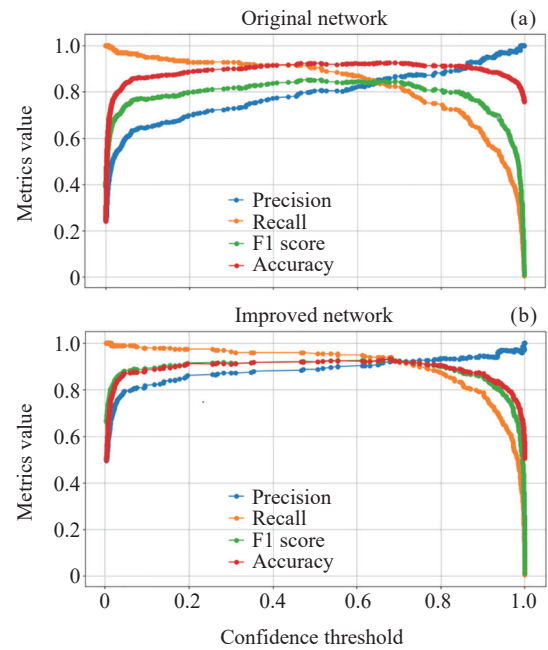


图 11 两种网络的置信度评价

Fig. 11 Confidence evaluation of the original network

4 结论

在 VGG16 基础上融入定制化自注意力机制的作用是通过突出关键特征和整合时频信息, 提高模型在识别任务中的性能, 这有助于模型在复杂和多样化的环境下, 更准确地识别出极光千米波.

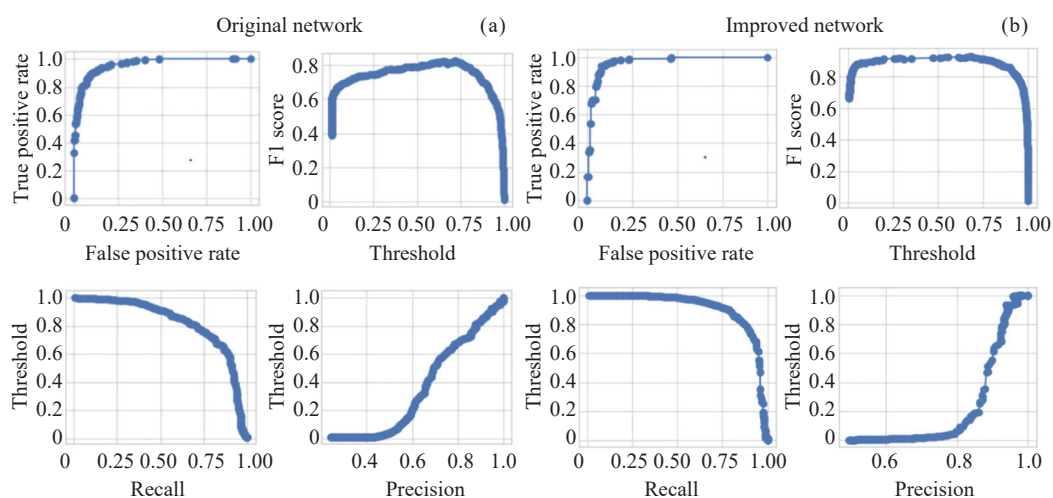


图 12 两种网络的性能指标

Fig. 12 Performance indice of the two networks

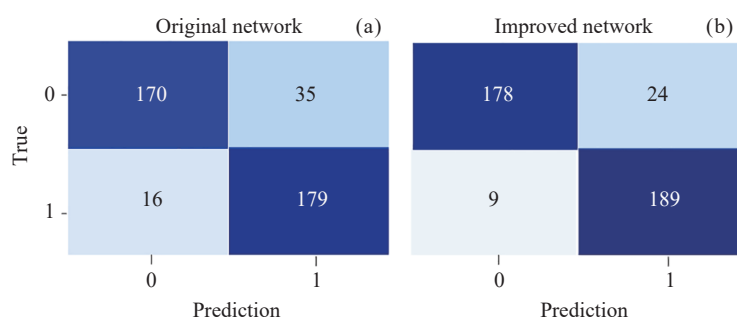


图 13 两种网络的混淆矩阵对比

Fig. 13 Comparison of the confusion matrices of the two networks

融入的过程可概括为: SAM-V 通过对特征图中每个位置与其他所有位置的关系进行建模, 得到注意力权重矩阵后再进行加权特征融合, 将融合后的特征输入全连接层进行分类, 同时特征映射得到每个类别的概率. 在原始 VGG16 网络的卷积层和池化层没有其他改变的情况下, 采用 ReLU 激活函数改善梯度分布. 同时, 采用基于 AKR 特征的线性学习率预热和动态衰减, 可以避免初始梯度消失或爆炸, 防止模型过拟合并保持稳定性. 同时采用 Adam 优化器更新网络参数, 使得模型训练更可控. 通过文中的评价指标可以观察到, 模型改进后的精确率、召回率、置信度评价以及 AUC 相对于原模型均有提升, 正负类预测率有所增加, 模型识别 AKR 性能增强. 同时, 部分易被误识别的数据也得到类别更正和概率提升.

未来可聚焦于实际遥感图像复杂情景下的各种因素影响, 研究极光千米波的所有场景特征, 增加多

方位实验并改进模型, 以达到更强的适应性和实用性.

参考文献

- [1] ZHAO Wanli. The Global Distribution Characteristics of Auroral Kilometric Radiations in the Radiation Belt[D]. Changsha: Changsha University of Science & Technology, 2019 (赵万里. 地球辐射带区域极光千米波的全球分布特征研究 [D]. 长沙: 长沙理工大学, 2019)
- [2] GURNETT D A. The Earth as a radio source: terrestrial kilometric radiation[J]. *Journal of Geophysical Research*, 1974, **79**(28): 4227-4238
- [3] KURTH W S, BAUMBACK M M, GURNETT D A. Direction-finding measurements of auroral kilometric radiation[J]. *Journal of Geophysical Research*, 1975, **80**(19): 2764-2770
- [4] PRITCHETT P L, WINGLEE R M. Generation and propagation of kilometric radiation in the auroral plasma

- cavity[J]. *Journal of Geophysical Research: Space Physics*, 1989, **94**(A1): 129-143
- [5] BENEDIKTOV E A, GETMANTSEV G G, SAZONOV Y A, *et al*. Preliminary results of measurements of the intensity of distributed extraterrestrial radio-frequency emission at 725 and 1525-kHz frequencies by the satellite electron-2[J]. *Kosmicheskie Issledovaniya*, 1965, **3**(6): 791-794
 - [6] CALVERT W. The auroral plasma cavity[J]. *Geophysical Research Letters*, 1981, **8**(8): 919-921
 - [7] LI Wentao. Research on the Influence of Outer Radiation Belt Electron Dynamic Evolution by Auroral Kilometric Radiation[D]. Changsha: Changsha University of Science & Technology, 2021 (李文涛. 极光千米波对外辐射带高能电子动力学演化过程的影响研究 [D]. 长沙: 长沙理工大学, 2021)
 - [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the 3rd International Conference on Learning Representations. San Diego: ICLR, 2015
 - [9] WANG Fei, YANG Qiuju. Classification of auroral images based on convolutional neural network[J]. *Chinese Journal of Polar Research*, 2018, **30**(2): 123-131 (王菲, 杨秋菊. 基于卷积神经网络的极光图像分类 [J]. 极地研究, 2018, **30**(2): 123-131)
 - [10] LI Yanzhi, CHEN Changhong, XIE Xiaofang. Aurora image classification algorithm based on improved convolution neural network[J]. *Journal of Nanjing University of Posts and Telecommunications: Natural Science*, 2019, **39**(6): 86-93 (李彦枝, 陈昌红, 谢晓芳. 基于改进卷积神经网络的极光图像分类算法研究 [J]. 南京邮电大学学报: 自然科学版, 2019, **39**(6): 86-93)
 - [11] NIU Chuang. Deep Learning Based Method for Auroral Local Structure Recognition and Localization[D]. Xi'an: Xidian University, 2020 (牛闯. 基于深度学习的极光局部结构识别与定位 [D]. 西安: 西安电子科技大学, 2020)
 - [12] REN Jie. Automatic Recognition and Detection of Aurora Events Based on Deep Learning[D]. Xi'an: Shaanxi Normal University, 2022 (任杰. 基于深度学习的极光事件自动识别与检测 [D]. 西安: 陕西师范大学, 2022)
 - [13] VASWANI A, SHAZEER N, PARMAR N, *et al*. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 6000-6010
 - [14] YUE Hongguang, HAN Longmei, WANG Zhengyong, *et al*. Remote sensing image scene classification based on multi-channel self-attention network[J]. *Journal of Sichuan University: Natural Science Edition*, 2023, **60**(2): 023002 (岳泓光, 韩龙玫, 王正勇, 等. 基于多通道自注意力网络的遥感图像场景分类 [J]. 四川大学学报: 自然科学版, 2023, **60**(2): 023002)
 - [15] CHANG Xilong, LIANG Kun, LI Wentao. Review of development of deep learning optimizer[J]. *Computer Engineering and Applications*, 2024, **60**(7): 1-12 (常禧龙, 梁琨, 李文涛. 深度学习优化器进展综述 [J]. 计算机工程与应用, 2024, **60**(7): 1-12)
 - [16] PAN Meiyang, CAI Xingyu, XUE Jian. An aerial target recognition method for radar based on time-frequency spectrogram and CNN[J]. *Fire Control Radar Technology*, 2023, **52**(4): 16-22 (潘美艳, 蔡兴雨, 薛健. 基于时频谱图和 CNN 的雷达空中目标识别方法 [J]. 火控雷达技术, 2023, **52**(4): 16-22)
 - [17] HUAWEI TECHNOLOGIES CO., LTD. A Channel Prediction Method and Related Equipment: CN201811054284.7[P]. 2018-09-08 (华为技术有限公司. 一种信道预测方法及相关设备: CN 201811054284.7[P]. 2018-09-08)
 - [18] ZHANG Liang. AdamX Optimizer: A New Optimizer Based on Gradient and Momentum Coordination of Learning Rate[D]. Ji'nan: Shandong University, 2022 (张良. AdamX 优化器: 基于梯度和动量协调控制学习率的一种新型优化器 [D]. 济南: 山东大学, 2022)
 - [19] ZHU Shengwei, DENG Xiaojian, ZHANG Guangzhi. Lithology Classification of Earthquakes Based on Deep Neural Networks[C]//Proceedings of the 2018 Joint Academic Annual Meeting of Chinese Earth Sciences. Beijing: Chinese Geophysical Society, China Seismological Society, Organizing Committee of the National Symposium on Petrology and Geodynamics, Professional Committee on Tectonic Geology and Geodynamics of the Chinese Geological Society, Professional Committee on Regional Geology and Mineralization of the Chinese Geological Society, Earth Sciences Division of the National Natural Science Foundation of China, 2018 (朱圣伟, 邓小江, 张广智. 基于深度神经网络的地震岩性分类 [C]//2018 年中国地球科学联合学术年会论文集. 北京: 中国地球物理学会, 中国地震学会, 全国岩石学与地球动力学研讨会组委会, 中国地质学会构造地质学与地球动力学专业委员会, 中国地质学会区域地质与成矿专业委员会, 国家自然科学基金委员会地球科学部, 2018)
 - [20] JIANG Rui, SUN Liuting, WANG Xiaoming, *et al*. Research on EEG signal classification of motor imagery based on AE and transformer[J]. *Chinese Journal on Internet of Things*, 2023, **7**(1): 118-128 (蒋锐, 孙刘婷, 王小明, 等. 基于 AE 和 Transformer 的运动想象脑电信号分类研究 [J]. 物联网学报, 2023, **7**(1): 118-128)
 - [21] MA Lukuan. Research on Pedestrian Reidentification Based on Data Enhancement and Attribute Assist[D]. Chengdu: Sichuan University, 2021 (马路宽. 基于数据增强和属性辅助的行人重识别研究 [D]. 成都: 四川大学, 2021)
 - [22] XIAO F L, TANG J W, ZHANG S, *et al*. Asymmetric distributions of auroral kilometric radiation in earth's northern and southern hemispheres observed by the Arase satellite[J]. *Geophysical Research Letters*, 2022, **49**(13):

e2022GL099571

- [23] WANG Huajian, XU Li, DIAO Zhicheng, *et al.* Audio processing method and device, electronic equipment and storage medium: CN202310997762.2[P]. 2023-08-07 (汪华健, 许丽, 刁志程, 等. 音频处理方法、装置、电子设备和存储介质: CN202310997762.2[P]. 2023-08-07)
- [24] ZHANG S, LIU S, LI W T, *et al.* A concise empirical formula for the field-aligned distribution of auroral kilometeric radiation based on Arase satellite and van Allen probes[J]. *Geophysical Research Letters*, 2021, **48**(8): e2021GL092805
- [25] ZHANG S, SHANG X J, HE Y H, *et al.* Dominant roles of high harmonics on interactions between AKR and radiation belt relativistic electrons[J]. *Geophysical Research Letters*, 2020, **47**(16): e2020GL088421
- [26] ZHANG S, YIN Q P, YANG H M, *et al.* Direct observation of l-x mode of auroral kilometeric radiation in the lower latitude magnetosphere by the Arase satellite[J]. *Geophysical Research Letters*, 2024, **51**(5): e2023GL105694
- [27] YAN Xiaochen. Research on Chemical Process Fault Diagnosis Method Based on Improved Resnet[D]. Beijing: Beijing University of Chemical Technology, 2023 (严孝宸. 基于改进残差神经网络的化工过程故障诊断方法研究 [D]. 北

京: 北京化工大学, 2023)

- [28] ZHAO Xuejiao, JIANG Yong, SU Luqing. Research on image recognition technology based on deep learning algorithm[J]. *Modeling and Simulation*, 2023, **12**(3): 3024-3034 (赵雪娇, 江湧, 苏露情. 基于深度学习算法的图像识别技术研究 [J]. *建模与仿真*, 2023, **12**(3): 3024-3034)

作者简介



王鹤野 女, 2003年4月出生于山东省滨州市, 现为长沙理工大学物理与电子科学学院大四学生, 主要参与机器人运控, 机器视觉 CV 等应用与研究.
E-mail: 3163797415@qq.com



张赛 男, 1986年10月出生于山西省运城市, 现为长沙理工大学物理与电子科学学院副教授, 硕士生导师, 主要研究方向为空间等离子体物理和人工智能应用等.
E-mail: saizh@126.com