# 基于双域注意力增强融合网络的日冕物质抛射图像匹配优化算法

郭媛[1,2,3] 蔡燕霞[1,3] 闫帅楠[1,2,3] 包黎莉[1,3,*]

1 (中国科学院国家空间科学中心，北京 100190)

2 (中国科学院大学，北京 100049)

3 (中国科学院太阳活动与空间天气全国重点实验室，北京 100190)

**摘要** 日冕物质抛射（Coronal Mass Ejection, CME）是引发灾害性空间天气事件的重要因素，对其传播过程的动态可视化有助于理解相关物理机制并支持预警决策。增强现实（Augmented Reality, AR）技术能够直观展示 CME 传播特性，但其应用依赖于高精度的 CME 图像匹配。针对现有图像匹配方法特征表达能力不足、易受噪声干扰等问题，本文提出一种基于双域注意力增强融合网络（Dual-Domain Attention Enhanced Fusion Network, DDFN）的图像匹配算法。该算法在卷积神经网络和视觉 Transformer 基础上引入空间域和通道域注意力机制强化特征表达，并利用对比学习框架提升模型对噪声和亮度变化的鲁棒性。结果表明，DDFN 在原始及噪声数据集上的 Top-1 匹配准确率分别达到 81.76%和 68.44%，优于对比模型。该算法已部署于 AR 系统，实现了 CME 图像的高效匹配与动态可视化。

**关键词** 空间天气；日冕物质抛射；对比学习；双域注意力；图像匹配

**中图分类号** P353

最新录用

# An Optimization Algorithm for Coronal Mass Ejection Image Matching Based on Dual-Domain Attention Enhanced Fusion Network

GUO Yuan[1,2,3] CAI Yanxia[1,3] YAN Shuainan[1,2,3] BAO Lili[1,3,*]

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

2 (*University of Chinese Academy of Sciences, Beijing* 100049)[1]

3 (*State Key Laboratory of Solar Activity and Space Weather, Chinese Academy of Sciences, Beijing* 100190)

**Abstract** Coronal mass ejections (CMEs) are one of the major drivers of severe space weather events. Dynamic visualization of their propagation process is important for understanding the underlying physical mechanisms and for supporting space weather monitoring and warning decisions. Augmented reality (AR) technology provides an intuitive way to present CME propagation characteristics by integrating virtual and real scenes. However, the effective application of AR in this context relies on high-precision CME image matching, which remains challenging because existing image-matching methods often suffer from insufficient feature representation and are highly sensitive to noise. To address these problems, this paper proposes a CME image matching algorithm based on a Dual-Domain Attention Enhanced Fusion Network (DDFN). The proposed method introduces spatial-domain and channel-domain attention mechanisms on the basis of convolutional neural networks and vision transformer architectures, thereby enhancing feature representation from complementary perspectives. Specifically, the convolutional structure is used to capture local image patterns, while the Vision Transformer contributes to global contextual modeling. The dual-domain attention design further strengthens discriminative feature extraction by emphasizing informative spatial regions and important feature channels. In addition, a contrastive learning framework is employed to improve the robustness of the model against noise interference and brightness variations, which are common in CME image data and may degrade matching performance. Experimental results demonstrate that the proposed DDFN achieves Top-1 matching accuracies of 81.76% and 68.44% on the original and noisy datasets, respectively, outperforming the comparison models. These results indicate that the proposed algorithm can effectively improve CME image matching accuracy while maintaining robustness under degraded imaging conditions. Furthermore, the algorithm has been deployed in an AR system, enabling efficient CME image matching and dynamic visualization of the CME propagation process. The proposed method provides a technical basis for integrating intelligent image matching with AR-based space weather visualization, and it offers potential support for the analysis, interpretation, and visual presentation of CME propagation behavior.

**Key words** Space weather; Coronal Mass Ejection; Contrast learning; Dual-Domain Attention; Image matching

## 0 引言

空间天气是指由太阳活动及其与地球空间环境相互作用所产生的一系列现象。其中，日冕物质抛射（Coronal Mass Ejection, CME）作为重要的太阳活动，能引发灾害性空间天气事件，严重影响航天器、电力系统及通信网络等关键基础设施[1-3]。近年来，CME 成为空

[1]* 基金项目：中国科学院重点部署项目（ZDRE-KT-2021-3）
投稿日期：2025 年 4 月 29 日
E-mail: guoyuan22@mails.ucas.ac.cn

间天气研究和航天应用中的重点对象[4]，其传播过程的动态可视化有助于深入理解传播机制并提升灾害预警能力。

传统 CME 可视化多采用二维动态图像或三维静态模型，难以有效展现其时空演化过程。增强现实（Augmented Reality, AR）技术通过将虚拟信息叠加于真实环境，提供沉浸式交互体验，已在医疗、教育等领域广泛应用[5-12]，但在空间天气领域中尚处于起步阶段。跟踪注册技术是实现 AR 技术的关键技术，其中基于自然特征的跟踪注册技术依赖于高效准确的图像特征提取和匹配算法。由于 CME 图像结构稀疏、边界模糊且噪声复杂，同时不同事件视觉相似，给特征提取和匹配带来了挑战。

图像匹配通常包含两个关键环节：特征提取和特征匹配[13]。特征提取负责将图像转换为判别性强的特征表示，而特征匹配则通过度量特征相似性建立图像间的对应关系。在特征匹配方面，研究者提出了多种创新方法提升匹配精度。例如，图神经网络[14]通过捕捉节点间的复杂关系，提升匹配一致性；生成式对抗网络[15]被引入匹配任务，通过对抗训练增强跨域特征对齐能力。然而，这类方法网络结构复杂、计算量大[16]，在资源受限的 CME 实时匹配任务中应用受限。一般而言，匹配的准确率主要受限于特征提取阶段的表达能力，对于 CME 图像匹配任务，获取高质量的特征表示往往比采用复杂匹配算法更为关键和高效。因此，本文聚焦于特征提取阶段的优化。常用的图像特征提取方法如 SIFT[17]、ORB[18]、SURF[19]和 BRISK[20]等，能够在一定程度上应对图像的旋转、缩放等几何变换[21]，但在复杂结构和模糊背景干扰显著的 CME 图像匹配任务中，其性能仍受较大限制，难以满足高精度匹配的需求。近年来，深度学习展现出优越性能[22]。卷积神经网络（Convolutional Neural Network, CNN）如 AlexNet[23]、ResNet[24]和 VGG[25]等被广泛应用于太阳黑子分类[26-27]、耀斑预报[28-29]及 CME 检测[30-31]，但 CNN 对长距离依赖建模能力有限，难以捕捉 CME 整体形态和传播路径的全局信息。视觉 Transformer（Vision Transformer, ViT）利用自注意力机制增强全局信息建模，部分弥补了 CNN 的局限[32-33]，但对局部细节敏感度较低，且依赖大规模训练数据，难以适应数据稀缺的 CME 领域。为此，混合 CNN 与 Transformer 模型被提出，结合局部与全局特征优势，在太阳观测任务中取得一定成效[34-35]，但在特征表达平衡与噪声抑制方面仍存在挑战。注意力机制被证实可有效提升特征表达能力，并已在太阳磁图超分辨率和黑子检测中应用[36-37]。然而，融合模型和注意力机制在 CME 图像特征提取中的系统研究仍较缺乏，亟需深入探索。

针对上述问题，本文提出基于双域注意力增强融合网络（Dual-Domain Attention

最新录用

Enhanced Fusion Network, DDFN）的 CME 图像匹配优化算法。该方法融合 CNN 的局部特征提取与 ViT 的全局建模能力，引入空间域与通道域注意力机制以增强特征表达的判别性，并结合对比学习框架 Simple Contrastive Learning of Representations（SimCLR）提升特征对齐与泛化能力，实现高鲁棒性和精度的 CME 图像匹配，助力空间天气研究与 AR 系统应用。

## 1 数据与预处理方法

### 1.1 数据源

CME 图像数据主要来源于 SOHO 卫星大角度和光谱测量日冕仪（Large Angle and Spectrometric Coronagraph, LASCO）[38]C2 子仪器拍摄的白光日冕图像。该仪器具备高动态范围成像能力，能有效捕捉日冕层大尺度结构变化。

考虑到现有公开 CME 与现实场景匹配数据集匮乏，本文构建了包含 847 幅 CME 观测图像和 847 幅现实场景图像的新数据集，示例如图 1 所示。CME 图像涵盖 CDAW 记录的 2000 至 2022 年间的多个典型事件，涉及速度约 500 至 2000 km/s 以上，多种形态如全晕、偏晕和窄带，涵盖日面、边缘和背面的不同起源位置，以及爆发、扩张和传播等不同演化阶段。现实场景图像分辨率统一为 3072×4096，采集过程特别考虑了光学畸变、镜头变形等问题，通过多样化拍摄条件模拟 AR 系统实际环境需求。




(a) CME 观测图像数据集示例　　(b) CME 现实场景图像数据集示例

(a) Example of the CME observation image dataset　(b) Example of the CME real-world scene image dataset

图 1　CME 图像数据集示例

Fig. 1 Examples of CME image datasets

### 1.2 基于显著性特征的 CME 图像自适应增强方法

CME 在观测图像中存在显著差异，部分事件边界清晰、对比度高，另一些则微弱模糊，需自适应处理策略。传统统一增强方法无法兼顾不同显著性区域，容易导致过度或不足增强。为此，本文提出基于显著性特征的自适应增强方法（Adaptive Enhancement Method

最新录用

Based on Significant Features, AEM）。该方法首先通过计算局部对比度和边缘强度特征，评估图像中不同区域的显著性程度。随后根据显著性评估结果，自适应选择增强策略和参数。对于高显著性区域，采用温和增强以保持原有特征；对于低显著性区域，则采用较强增强以突出潜在特征。此策略有效克服了传统方法的局限，为后续 CME 图像匹配提供更优质的图像基础。

显著性评估通过局部对比度与边缘强度的加权计算实现。对于图像像素点$(x,y)$，其局部对比度特征为：

$$C(x,y)=\frac{|I(x,y)-\mu_L|}{\sigma_L+\epsilon} \tag{1}$$

其中，$I(x,y)$为像素点灰度值，$\mu_L$和$\sigma_L$分别为局部窗口的均值和标准差，$\epsilon$用于防止分母为零。边缘强度$E(x,y)$基于 Sobel 算子梯度幅值获得：

$$E(x,y)=\sqrt{\left(\frac{\partial I}{\partial x}\right)^2+\left(\frac{\partial I}{\partial y}\right)^2} \tag{2}$$

其中，$\frac{\partial I}{\partial x}$ 和$\frac{\partial I}{\partial y}$分别为图像$x$、$y$方向的梯度，以捕获 CME 结构边界，有助于区分图像结构与噪声。基于以上特征，显著性图$S$在点$(x,y)$处的特征为：

$$S(x,y)=w_1C(x,y)+w_2E(x,y) \tag{3}$$

其中，$w_1$，$w_2$为权重系数，用于衡量对比度和边缘信息对特征的贡献程度。显著性图的统计特性定义如下：

$$\mu_S=\frac{1}{mn}\sum_{x=1}^{m}\sum_{y=1}^{n}S(x,y) \tag{4}$$

$$\sigma_S=\sqrt{\frac{1}{mn}\sum_{x=1}^{m}\sum_{y=1}^{n}(S(x,y)-\mu_S)^2} \tag{5}$$

其中，$\mu_S$和$\sigma_S$分别为显著图的全局均值和标准差，m 和 n 分别代表图像的高度和宽度。显著性阈值通过自适应方式确定：

$$T_{high}=\mu_S+\alpha\sigma_S \tag{6}$$

$$T_{low}=\mu_S-\alpha\sigma_S \tag{7}$$

其中，α 为调节参数。阈值$T_{high}$和$T_{low}$将图像划分为高、中、低三个显著性区域，并采用

最新录用

差异化增强策略来获得不同区域的增强像素值$I_{AEM}(x,y)$：

$$I_{AEM}(x,y)=\begin{cases} I(x,y)+\beta_1\left(I(x,y)-\mu_L\right), S(x,y)\geq T_{high} \\ I(x,y)+\beta_2\left(I(x,y)-\mu_L\right)+\gamma_1\nabla^2 I(x,y), T_{low}\leq S(x,y)<T_{high} \\ I(x,y)+\beta_3\left(I(x,y)-\mu_L\right)+\gamma_2\nabla^2 I(x,y)+\delta_1 H\left(I(x,y)\right), S(x,y)<T_{low} \end{cases} \tag{8}$$

其中，$\beta_1$、$\beta_2$、$\beta_3$表示对比度增强系数，$\gamma_1$、$\gamma_2$为拉普拉斯算子增强系数，用以提升CME结构的可见性与边缘清晰度；$\delta_1$为自适应直方图均衡化权重，以抑制增强不足或过度。$\nabla^2$表示拉普拉斯算子，用于细节增强。$H(\cdot)$为自适应直方图均衡化操作。

为消除不同区域间视觉上的不连贯性，对区域边界处采用平滑过渡机制，如式(9)所示：

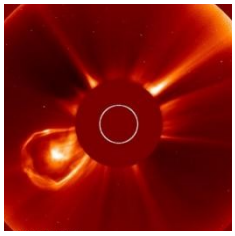
$$W=\frac{1}{1+e^{-k|S-T|}} \tag{9}$$

其中，$k$控制过渡陡峭程度，$T$表示过渡阈值。最终增强后的图像$I_e$表示为：

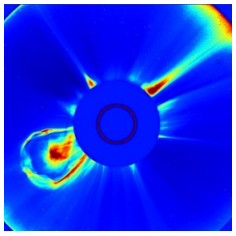
$$I_e=W I_{AEM}+(1-W)I_o \tag{10}$$

其中，$I_o$为原始图像。

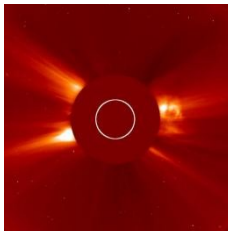
本文选取LASCO C2观测图像中具有代表性的CME事件进行实验分析。实验数据包括2022年9月26日（结构特征明显的CME）和2016年7月23日（结构特征微弱的CME）的原始观测图像和现实场景图像。实验中，网络的输入图像大小统一预处理为256×256，图 2展示了增强前后的对比效果，增强后的结果以热力图表示以清晰展示增强差异。从结果来看，对于结构特征明显的CME事件（图 2(a)-(b)、(e)-(f)），AEM有效突出CME结构及边界，改善了原始图像中边缘亮度饱和和内部模糊问题，同时抑制了设备拍摄引起的噪声干扰，提高了特征辨识度。对于结构较弱的CME事件（图 2(c)-(d)、(g)-(h)），采用复合增强策略显著提升了低显著性区域的对比度，增强了暗弱等离子体团的可见性，使磁力线拓扑结构更清晰。实验表明，所提方法在保护强结构完整性的同时，能有效挖掘微弱特征。



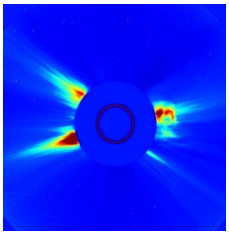
(a) 2022年9月26日CME图像增强前
(a) CME image before



(b) 2022年9月26日CME图像增强后
(b) CME image after

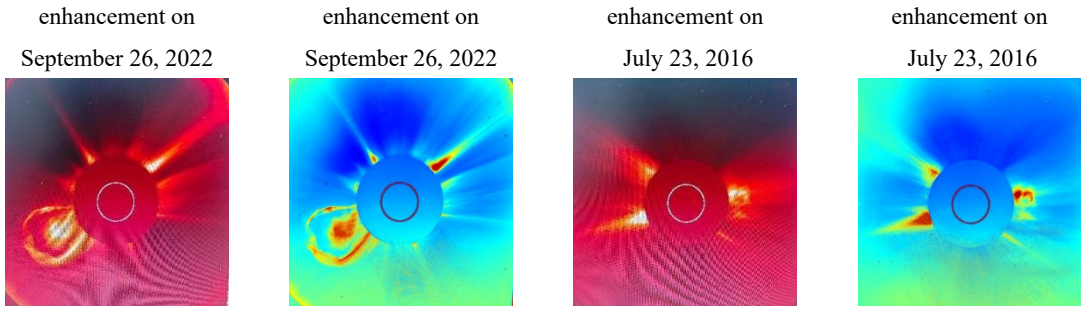


(c) 2016年7月23日CME图像增强前
(c) CME image before



(d) 2016年7月23日CME图像增强后
(d) CME image after

最新录用




(e) 2022 年 9 月 26 日现实场景图像增强前 (e) Real-world scene image before enhancement on September 26, 2022

(f) 2022 年 9 月 26 日现实场景图像增强后 (f) Real-world scene image after enhancement on September 26, 2022

(g) 2016 年 7 月 23 日现实场景图像增强前 (g) Real-world scene image before enhancement on July 23, 2016

(h) 2016 年 7 月 23 日现实场景图像增强后 (h) Real-world scene image after enhancement on July 23, 2016

图 2　CME 图像自适应增强预处理结果

Fig. 2 The adaptive enhancement preprocessing results of CME images

## 2 基于双域注意力增强融合网络的 CME 图像匹配优化算法

### 2.1 双域注意力增强融合网络

为有效融合局部细节与全局语义信息，提升 CME 图像的表征能力，本文提出了 DDFN 模型。首先，输入的 CME 图像经过 AEM 预处理，以突出显著特征。随后，DDFN 结合了多尺度 CNN 与 ViT 两种特征提取方法，在 CNN 提取的特征基础上，显式引入通道注意力机制（Channel Attention Mechanism, CAM）和空间注意力机制（Spatial Attention Mechanism, SAM），进一步增强特征表达能力。最终，将双域注意力增强的局部特征与 ViT 提取的全局特征融合，并通过多层感知机（Multilayer Perceptron, MLP）进行降维，实现高效的特征整合（如图 3 所示）。

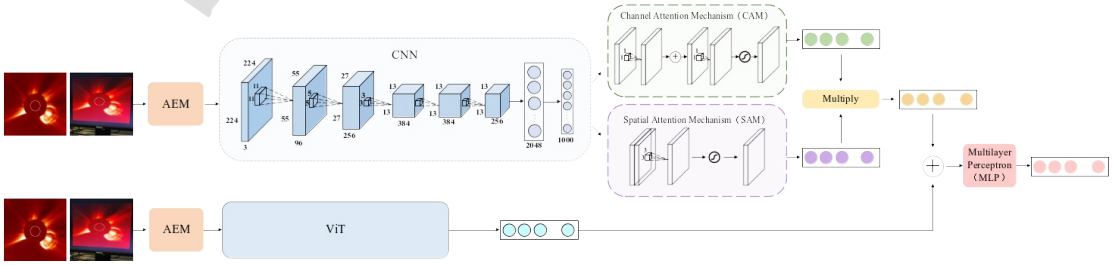



图 3　DDFN 网络结构

Fig. 3 The structure of Dual-Domain Attention Enhanced Fusion Network

针对 CME 图像的空间复杂性及尺度多变性，采用 5 层多尺度卷积结构依次提取特征：

$$F_{CNN}^{(l)} = ConvBlock_l\left(F_{CNN}^{(l-1)}\right), l \in \{1,2,3,4,5\} \tag{11}$$

其中，$ConvBlock_l$为卷积模块。第一层使用 11×11 卷积核捕获大尺度信息，后续层使用小

卷积核提取细节。第 5 层输出经展平得到局部特征：

$$F_{CNN}=Flatten\left(F_{CNN}^{(5)}\right)\in R^{256\times 6\times 6} \tag{12}$$

基于最后一层 CNN 特征图，引入通道注意力和空间注意力，进一步提升特征表达能力。通道注意力通过全局平均池化后，利用两层全连接计算注意力权重，捕捉 CME 图像通道的重要性程度：

$$W_c=\sigma\left(W_2\cdot\delta\left(W_1\cdot GAP\left(F_{CNN}^{(5)}\right)\right)\right) \tag{13}$$

其中，GAP(·)表示全局平均池化。空间注意力通过最大池化与平均池化拼接后卷积，实现对空间位置的重要性建模：

$$W_s=\sigma\left(Conv\left(\left[MaxPool\left(F_{CNN}^{(5)}\right);AvgPool\left(F_{CNN}^{(5)}\right)\right]\right)\right) \tag{14}$$

其中，$\sigma$和$\delta$分别为 Sigmoid 和 ReLU 激活函数。增强后的特征表达如下：

$$F_{CNN}^{att}=F_{CNN}^{(5)}W_cW_s \tag{15}$$

为获取图像全局信息，DDFN 将 CME 图像 X 输入 ViT 模型，通过多层 Transformer 编码器及线性映射，得到全局特征向量：

$$F_{ViT}=ViT(X)\in R^{1000} \tag{16}$$

之后，将双域注意力增强的 CNN 局部特征$F_{CNN}^{att}$与 ViT 全局特征拼接，并通过 MLP 对融合特征进行维度压缩，得到最终的特征向量：

$$F_{fused}=Concatenate\left[F_{CNN}^{att};F_{ViT}\right] \tag{17}$$

$$F_{out}=W_1\cdot F_{fused}+b_1 \tag{18}$$

其中，$W_1\in R^{|9216+1000|\times 512}$，$b_1\in R^{512}$，最终输出$F_{out}\in R^{512}$可作为后续网络结构的入口表征。

### 2.2 模型整体架构

基于前节提出的 DDFN，本文设计了完整的 CME 图像匹配优化模型架构，以提升模型对数据质量差异的适应能力及匹配精度。整体模型框架如图 4 所示。

该模型首先利用 DDFN 分别对输入的 CME 图像进行局部多尺度特征提取和全局信息捕获，生成具有丰富语义和空间信息的表征。随后，将提取的特征输入对比学习框架 SimCLR[39]，通过构建正负样本对，实现特征的对齐与差异性学习，增强模型对特征细微变

最新录用

化的敏感度与判别能力。最终，经过对比学习优化的特征用于指导 CME 图像匹配任务，提高匹配的准确性和鲁棒性。

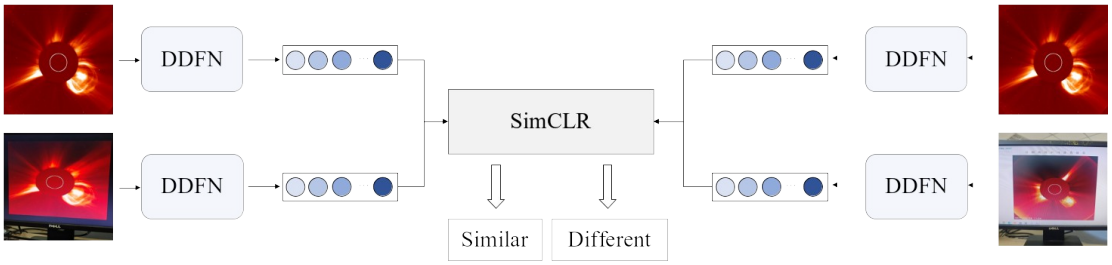



图 4　模型架构

Fig. 4 Model architecture

## 3 实验结果与分析

### 3.1 评估指标

（1）损失函数

本文以 SimCLR 为对比学习的基础框架，其核心为对比损失函数。该函数通过最大化正样本对之间的相似性，同时最小化负样本对之间的相似性，实现学习具有判别性的特征表示。具体地，本文采用 NT-Xent Loss（Normalized Temperature-Scaled Cross Entropy Loss）作为对比损失函数。对于一个批次的输入图像，首先通过数据增强生成视图，并计算它们之间的相似度得分。通过最小化 NT-Xent Loss，模型能够学习判别性特征表示，从而提升图像匹配的准确性。对于每一对正样本对，其损失函数定义为：

$$L_{NT-Xent}=-\log\frac{\exp\left(\frac{¿\left(z_i,z_j\right)}{\tau}\right)}{\sum_{k-1}^{2N} l_{[k\neq i]}\exp\left(\frac{¿\left(z_i,z_k\right)}{\tau}\right)} \tag{19}$$

其中，$¿\left(z_i,z_j\right)$为余弦相似度函数；$l_{[k\neq i]}$为指示函数，当$k\neq i$时取值为 1，否则为 0。$\tau$为温度参数，用于调节相似度分布。较小的温度参数会放大相似度得分的差异，使得模型更加关注高相似度的正样本对；而较大的温度参数则会平滑相似度得分的分布，有助于模型利用负样本信息。本文将设置为可学习参数，通过反向传播自动调整，以适应不同训练阶段的需求。

（2）实验环境和参数

本文实验在 Windows 11 操作系统上进行，采用 NVIDIA GeForce RTX 4060 GPU 加速。训练参数配置如表 1 所示。实验中的现实场景图像采用 Xiaomi 14 智能手机拍摄，曝光时间

最新录用

为 1/50~1/500s，ISO 范围为 100~500。

表 1　参数配置信息

Table 1 Parameter configuration information

| 参数配置项 | 值或名称 |
|---|---|
| SimCLR 温度参数$\tau$ | 0.1 |
| 学习率 | $1\times10^{-4}$ |
| 权重衰减 | $1\times10^{-2}$ |
| 训练轮次 | 100 |
| 优化器 | AdamW |

（3）性能评估指标

本文采用 Top-K 匹配准确率（Top-K Accuracy）评价模型性能，以衡量模型匹配结果的准确性，计算公式如式(20)所示：

$$Top_K\ Accuracy=\frac{1}{N}\sum_{i=1}^{N} I\left(y_i\in\left\{\hat{y}_i^{(1)},\hat{y}_i^{(2)},\ldots,\hat{y}_i^{(K)}\right\}\right) \tag{20}$$

其中，N 为匹配总数，$y_i$为第$i$张图像的真实目标图像，$\hat{y}_i^{(k)}$为模型对第 $i$ 张图像预测的第 $k$ 个最可能的目标图像，$I(\cdot)$为指示函数，满足条件时取值为 1，否则为 0。Top-1 Accuracy 表示匹配结果中预测的第 1 名是否为真实目标，Top-3 Accuracy 及 Top-5 Accuracy 分别表示真实目标是否包含于前 3 或前 5 名匹配结果内。

### 3.2 实验结果

为保障模型训练的有效性与泛化能力，CME 图像数据集按照 8：2 的比例划分为训练集与测试集，确保样本在事件类型、图像质量、结构复杂度等方面的代表性。为评估所提 DDFN 网络的特征表征能力，本文选取多种常规模型进行对比实验。

考虑到实际应用中图像因拍摄模糊、压缩失真及光照变化等因素产生的噪声干扰，本文设计了加入随机噪声的数据增强实验。实验具体采用噪声参数组合策略，即设置亮度、对比度和饱和度噪声参数为 0.2，色调噪声参数为 0.1，模拟真实场景下的不确定性影响，以测试模型对扰动信息的鲁棒性及容错能力，尤其关注特征提取与匹配阶段的稳定性。

表 2 展示了不同模型在原始数据集及加噪声数据集上的 Top-K 准确率。结果表明，DDFN 在各项指标中均表现优异，综合性能优于其他对比模型，显示了其在特征表达、鲁棒性及准确性方面的优势。具体而言，DDFN 在原始数据集上的 Top-1 准确率达 81.76%，较 CNN 和 ViT 模型分别提升了 8.23 和 9.67 个百分点，且优于 CNN 与 ViT 融合模型的 78.23%。Top-3 和 Top-5 准确率分别达到 95.29%和 97.65%，进一步验证了其稳定的特征提

最新录用

取能力。加入噪声后，所有模型性能均有所下降，但 DDFN 依然保持较高的准确率，Top-3 和 Top-5 准确率分别为 94.69%和 98.82%，在加噪声条件下仍表现出良好的抗干扰能力。性能下降方面，CNN、ViT 及融合模型的 Top-1 准确率分别降低了 14.83、16.28 和 15.11 个百分点，而 DDFN 下降幅度相对较小，仅为 13.32 个百分点，显示其对噪声干扰的更强容忍性。ViT 模型在全局注意力机制下，对噪声更敏感，导致部分性能下降，CNN 模型则表现出较强的鲁棒性，但在准确率提升上不及 DDFN。相比单一模型及浅层线性融合方案，DDFN 通过引入双域注意力机制并在双域特征间进行统一一维压缩融合，增强了深层语义信息整合能力。ViT 负责全局结构建模，CNN 补充局部细节，双域注意力模块强化关键特征表达，从而推动整体性能提升。本研究结果表明，DDFN 在两种数据集上均优于对比模型，验证了其特征融合策略的有效性与鲁棒性。

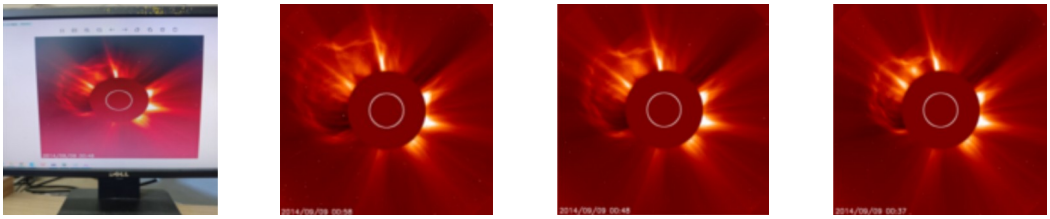
表 2　不同模型性能评估结果

Table 2 Performance evaluation results of different models in different dataset

| 模型 | 数据集 | Top-1 Accuracy | Top-3 Accuracy | Top-5 Accuracy |
|---|---|---|---|---|
| CNN | 原始数据集 | 73.53% | 88.82% | 91.76% |
| | 加噪声 | 58.70% | 85.54% | 92.33% |
| ViT | 原始数据集 | 72.09% | 93.02% | **97.67%** |
| | 加噪声 | 55.81% | 76.74% | 95.35% |
| CNN+ViT | 原始数据集 | 78.23% | 94.70% | 97.06% |
| | 加噪声 | 63.12% | 93.80% | 97.34% |
| **DDFN** | 原始数据集 | **81.76%** | **95.29%** | 97.65% |
| | 加噪声 | **68.44%** | **94.69%** | **98.82%** |

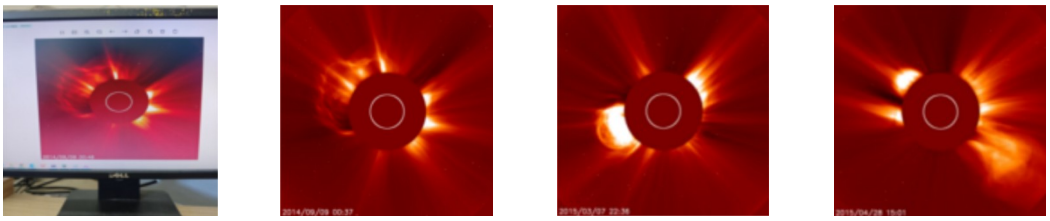
为直观展示各模型的匹配性能，本文选取具有明显抛射区域的 CME 现实场景图像，展示 DDFN 与 CNN、ViT 及其融合模型的匹配结果（见图 5）。图 5 依次展示了原始输入图像及 CNN、ViT、CNN+ViT 和 DDFN 在 Top-3 匹配结果上的表现。视觉对比显示，不同模型在特征表征和空间位置对齐上存在差异。CNN（图 5(a)）虽具一定纹理表达能力，但因局部感受野限制，抛射区域匹配存在不稳定和错配问题。ViT（图 5(b)）能够识别主体区域，但对边缘细节匹配不足。融合模型 CNN+ViT（图 5(c)）借助 Transformer 提升了全局感知能力，但特征融合欠完善，边缘细节仍有偏差。相比之下，DDFN（图 5(d)）在边缘对齐、区域响应及细节复原方面更为精准。其双域注意力机制通过通道维度突出关键滤波器，空间维度聚焦语义相关区域，有效抑制背景噪声，表现出较强的鲁棒性和泛化能力

录用稿



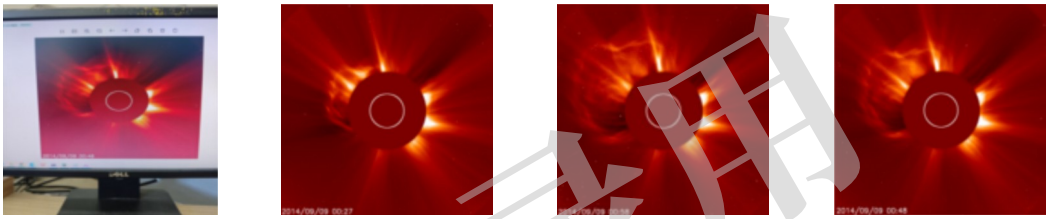
(a) CNN 匹配结果

(a) Results of CNN-based matching



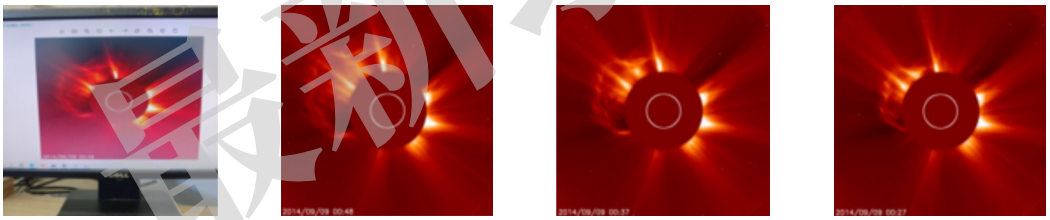
(b) ViT 匹配结果

(b) Results of ViT-based matching



(c) CNN+ViT 匹配结果

(c) Results of matching using CNN+ViT



(d) DDFN 匹配结果

(d) Results of DDFN-based matching

图 5　不同模型现实场景图像匹配结果

Fig. 5 Matching results of different models for real-world scene image

为了进一步了解 DDFN 的局限性并为后续改进提供方向，图 6 展示了 DDFN 模型的 Top-3 匹配结果错误示例。图中可见，现实场景图像中 CME 边缘轮廓微弱，伴有星尘杂散光，前两幅图像为同一 CME 事件中不同时刻的图像，第三幅图像为空间特征完全不一致的其他 CME 事件图像。这可能源于空间注意力在极低对比度下难以精准定位有效边缘，且通道注意力对弱信号的增强不足，致使特征相似度降低。错误案例的可视化分析表明，DDFN 在此类低对比度场景下仍存在不足，特别是在区分同一事件的图像时仍存在局限性。未来可通过数据增强和引入物理先验标签等手段，进一步提升模型鲁棒性。

最新录用

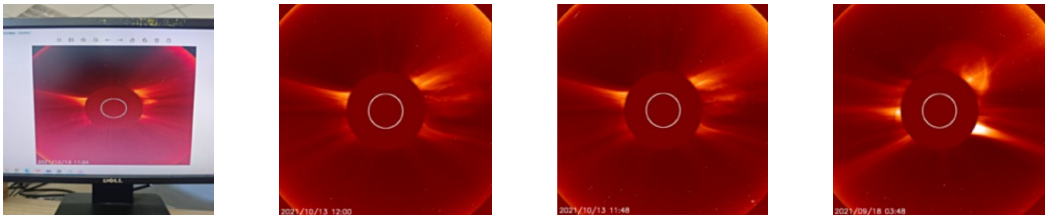
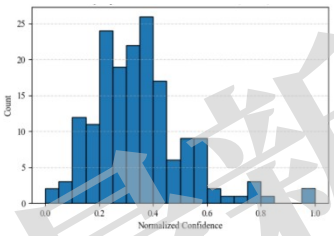


图 6　DDFN 现实场景图像误匹配结果

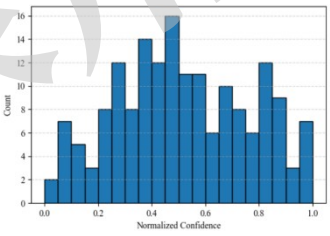
Fig. 6 Mismatch results of DDFN for real-world scene image

为深入评估模型在实际 AR 图像匹配任务中的表现，本文对不同模型基于 Top-3 匹配结果的置信度进行了归一化评估，置信度越接近 1 表示模型对匹配结果把握越高。图 7 展示了各模型的置信度分布直方图。CNN 模型置信度集中于 0.2 至 0.5，呈左偏分布，整体信心较低。ViT 模型分布较为均匀，覆盖广泛置信度区间，即 0.1 至 0.9。融合模型 CNN+ViT 的置信度较为平滑且集中，主要分布在 0.2 至 0.7 区间，但仍存在部分低置信度样本，表明特征融合存在信息损失。相比之下，DDFN 表现出更为紧凑且单峰的置信度分布，集中于 0.2 至 0.6 区间，显示出更稳定和可靠的匹配能力。结果表明 DDFN 通过双域注意力机制有效融合 CNN 与 ViT 特征，显著提升了匹配的可靠性和鲁棒性。



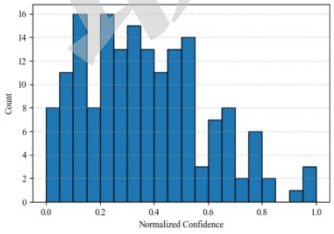

(a) CNN 置信度评估结果

(a) Confidence evaluation result of CNN



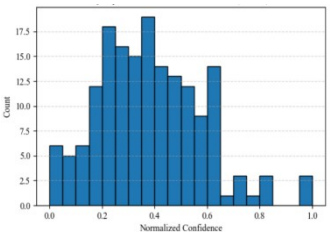

(b) ViT 置信度评估结果

(b) Confidence evaluation result of ViT




(c) CNN+ViT 置信度评估结果

(c) Confidence evaluation result of CNN+ViT




(d) DDFN 置信度评估结果

(d) Confidence evaluation result of DDFN

图 7　不同模型 Top-3 准确率置信度评估结果

Fig. 7 Confidence evaluation results of Top-3 Accuracy for different models

为探究双域注意力机制的贡献，本文在保持其余网络结构和训练设置不变的前提下，系统评估了空间域与通道域的独立及协同作用，评估结果见表 3。由表 3 可见，单独启用 SAM 时，模型在原始数据集的 Top-1 准确率为 80.59%，在加噪声数据集则下降至 64.89%，表明 SAM 虽有利于捕捉局部特征，但因依赖全局信息较强，抗噪能力有限。单

最新录用

独启用 CAM 时，模型在噪声数据集上的 Top-1 准确率提升至 66.67%，表现出更佳的稳定性，但细节特征提取仍存在不足。两者结合后，模型在两类数据集上均获得最高 Top-1 准确率，且加噪声数据集性能相较单独使用时有所提升，充分体现了双域注意力机制的协同效应及其对模型整体性能的促进作用。

表 3　注意力机制对模型整体性能影响的评估结果

Table 3 The evaluation results of the impact of the attention mechanism on the overall performance of the model

| SAM | CAM | 数据集 | Top-1 Accuracy | Top-3 Accuracy | Top-5 Accuracy |
|---|---|---|---|---|---|
| | √ | 原始数据集 | 79.41% | **95.88%** | 97.64% |
| | | 加噪声 | 66.67% | 94.98% | 98.52% |
| √ | | 原始数据集 | 80.59% | 94.12% | 97.05% |
| | | 加噪声 | 64.89% | **96.17%** | 97.94% |
| √ | √ | 原始数据集 | **81.76%** | 95.29% | **97.65%** |
| | | 加噪声 | **68.44%** | 94.69% | **98.82%** |

为进一步验证 AEM 模块的有效性，本文基于 DDFN 设计了消融实验。实验对比了未进行任何图像增强处理和 AEM 模块自适应增强处理的匹配准确率，实验结果如表 4 所示。由表 4 可见，引入 AEM 模块进行图像预处理的 DDFN 算法，在 Top-1 及 Top-5 匹配精度上均优于未采用图像预处理的算法。值得注意的是，在噪声干扰环境下，AEM 使 Top-5 准确率提升更为显著，表明该模块对低质量输入具有更强的适应性。实验结果进一步表明，AEM 模块能够根据 CME 图像不同区域的显著性差异，自适应调整增强策略，有效增强了图像特征的区分性与鲁棒性，提升了整体匹配性能。该结果验证了 AEM 方法在预处理阶段的有效性和实用性。虽然 DDFN 具备自主特征选择能力，但 AEM 作为预处理阶段的特征优化器，其价值主要体现在通过显著性引导的局部增强策略，有效提升图像关键区域的特征对比度，为后续特征提取提供了更优的输入条件。

表 4　AEM 模块对模型整体性能影响的评估结果

Table 4 The evaluation results of the impact of the AEM module on the overall performance of the model

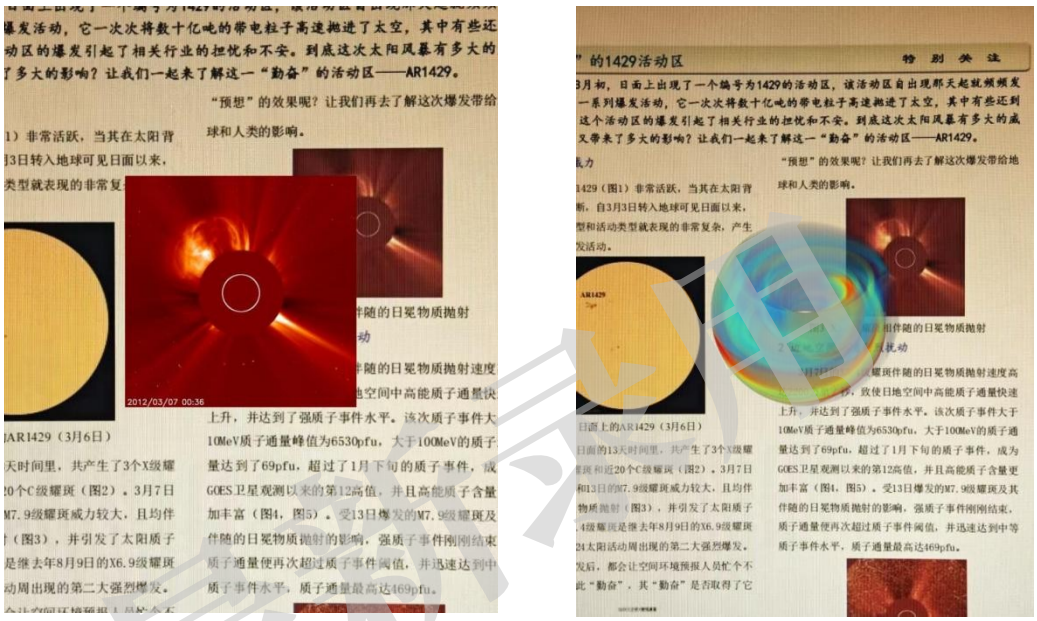
| AEM | 数据集 | Top-1 Accuracy | Top-3 Accuracy | Top-5 Accuracy |
|---|---|---|---|---|
| | 原始数据集 | 80.00% | **94.71%** | 96.47% |
| | 加噪声 | 66.96% | 94.10% | 97.93% |
| √ | 原始数据集 | **81.76%** | 95.29% | **97.65%** |
| | 加噪声 | **68.44%** | **94.69%** | **98.82%** |

## 4 算法应用

在本文中，CME 图像匹配算法旨在将包含 CME 的现实场景图像与数据库中的目标图像进行匹配，识别对应的 CME 事件。为验证 DDFN 算法的有效性，本文将其集成到 AR 系统中作为匹配模块。系统运行时，相机拍摄现实场景中的 CME 图像，匹配模块自动提取

最新录用

特征并检索数据库，找到对应的 CME 图像，随后加载该次事件的三维物理模型并叠加到实景中展示。三维 CME 模型基于社区协调建模中心（CCMC）提供的模拟数据，通过体绘制方法[40]生成，通过高效三维渲染技术，能够展现 CME 事件在行星际空间中的结构和演化过程。实验以 2012 年 3 月 7 日的 CME 事件为例（图 8），证明 DDFN 能够准确完成图像匹配任务并实现虚实融合，验证了 DDFN 在复杂场景下匹配能力的鲁棒性及工程适用性。相比传统特征提取方法，DDFN 在特征稀疏和噪声干扰条件下表现出更优的准确性与稳定性，有效缓解了 CME 图像误匹配问题。




(a) 现实场景图像匹配结果　　(b) AR 系统中 CME 可视化结果

(a) Matching result of real-world scene image　　(b) CME visualization result in AR System

图 8　AR 系统显示结果

Fig. 8 Display results of the AR system

此外，DDFN 还能应用于空间天气领域的 CME 相似事件推荐。通过学习当前 CME 事件的高维特征并结合 LASCO C2 的历史观测数据，算法可在历史数据库中精准匹配具有相似传播特性的事件，辅助预报员评估地磁暴强度和到达时间，提升预报的可靠性和决策效率。

## 5 结论

本文针对 CME 图像特征稀疏及匹配难题，提出了一种基于双域注意力机制的 DDFN 图像匹配优化算法，融合 CNN 与 ViT 特征，提升了特征判别性和鲁棒性。通过设计加噪声数据集验证算法在复杂环境下的泛化能力，实验结果表明 DDFN 在原始数据集和加噪声数据集上的 Top-1 匹配准确率分别达到 81.76%和 68.44%，优于多种对比模型。同时，结合

AEM 自适应增强预处理，进一步提升了匹配性能。

此外，算法已集成于 AR 系统并通过实际应用验证，显示出良好的实用性和工程潜力。尽管如此，模型在极端噪声和复杂背景下的鲁棒性仍有提升空间，未来工作将聚焦于优化模型结构、扩展数据集规模、探索更高效的注意力机制，并拓展算法在空间天气其他任务中的应用。

## 致谢

本文所使用的 CME 观测图像来源为 SOHO 卫星 LASCO C2 子仪器拍摄的白光日冕图像（https://cdaw.gsfc.nasa.gov/CME_list）。

## 参考文献

[1] DU Chenxi. Research in Ensemble Prediction of the Solar Energetic Particle Event[D]. Beijing: University of Chinese Academy of Sciences (National Space Science Center), 2022 (杜晨曦. 太阳高能粒子事件集合预报[D]. 北京: 中国科学院大学(中国科学院国家空间科学中心), 2022)

[2] LI Rong, QUAN Lin, LI Ling, *et al*. Review on status and development of space weather services[J]. *Progress in Geophysics*, 2023, **38**(6): 2417-2429 (李蓉, 全林, 李泠, 等. 空间天气服务现状及发展综述[J]. 地球物理学进展, 2023, **38**(6): 2417-2429)

[3] MIAO Juan, LI Zhitao, REN Tingling, *et al*. Statistical analysis on four kinds of space weather events during solar cycles 21~24[J]. *Chinese Journal of Space Science*, 2021, **41**(6): 849-857 (苗娟, 李志涛, 任廷领, 等. 第 21～24 太阳周 4 类空间天气事件爆发特征统计分析[J]. 空间科学学报, 2021, **41**(6): 849-857

[4] SHI Yurong. Study on Recommendation System for Coronal Mass Ejection Similar Events[D]. Beijing: University of Chinese Academy of Sciences (National Space Science Center), 2021 (石育榕. 日冕物质抛射相似事件推荐系统的研究[D]. 北京: 中国科学院大学(中国科学院国家空间科学中心), 2021)

[5] VAN BERGEM J S, VAN DOORN P F, HOOGENDOORN E M, *et al*. Gait and balance assessments with augmented reality glasses in people with Parkinson's disease: concurrent validity and test-retest reliability[J]. *Sensors*, 2024, **24**(17): 5485

[6] KAVIYARAJ R, UMA M. A survey on future of augmented reality with AI in education[C]//Proceedings of the 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS). Coimbator: IEEE, 2021: 47-52

[7] CHOI U, CHOI B. The effect of augmented reality on consumer learning for search and

最新录用

experience products in mobile commerce[J]. *Cyberpsychology*, *Behavior*, *and Social Networking*, 2020, **23**(11): 800-805

[8] KUHAIL M A, ABDULKERIM A Z, THORNQUIST E, *et al*. A review on the use of immersive technology in space research[J]. *Telematics and Informatics Reports*, 2025, **17**: 100191

[9] MEI Yingjie, WANG Dawei, SUN Chuanzhi, *et al*. A digital twin testing and adjusting system for aero-engine casings based on augmented reality[J]. *Acta Aeronautica et Astronautica Sinica*, 2024, **45**(21): 629462 (梅英杰, 王大伟, 孙传智, 等. 基于增强现实的航空发动机机匣数字孪生测调系统[J]. 航空学报, 2024, **45**(21): 629462)

[10] LI Jinfang, SU Jiancong, XIAO Libao, *et al*. Design of an AR visualization system for industrial robot teaching and training platform[J]. *Experimental Technology and Management*, 2024, **41**(4): 118-124 (李晋芳, 苏键聪, 肖立宝, 等. 工业机器人教学实训平台的 AR 可视化应用设计[J]. 实验技术与管理, 2024, **41**(4): 118-124)

[11] WANG Xin. Research on the Digital Display Design of the "Nanjing Museum Tang Dynasty Female Figure" Based on Augmented Reality Technology[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2023 (汪鑫. 基于增强现实技术的"南京博物院唐仕女俑"数字展示设计研究[D]. 南京: 南京邮电大学, 2023)

[12] BAN Rui, BING Yaxing, LI Yanqiang, *et al*. Design and application of a mobile AR museum artefact visualization system based on SURF[J]. *Qinghai Science and Technology*, 2023, **30**(5): 182-187 (班睿, 邴雅星, 李延强, 等. 基于 SURF 的移动 AR 博物馆文物可视化系统设计与应用[J]. 青海科技, 2023, **30**(5): 182-187)

[13] HUANG Yanyan. Image Matching Algorithms Based on Deep Learning[D]. Nanjing: Southeast University, 2023 (黄妍妍. 基于深度学习的图像匹配算法研究[D]. 南京: 东南大学, 2023)

[14] SCARSELLI F, GORI M, TSOI A C, *et al*. The graph neural network model[J]. *IEEE Transactions on Neural Networks*, 2009, **20**(1): 61-80

[15] MIRZA M, OSINDERO S. Conditional generative adversarial nets[OL]. arXiv preprint arXiv: 1411.1784, 2014

[16] LIU Haiqiao, LIU Meng, GONG Zichao, *et al*. A review of image matching methods based on deep learning[J]. *Acta Aeronautica et Astronautica Sinica*, 2024, **45**(3): 28796 (刘海桥, 刘萌, 龚子超, 等. 基于深度学习的图像匹配方法综述[J]. 航空学报, 2024, **45**(3): 28796)

[17] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, **60**(2): 91-110

[18] RUBLEE E, RABAUD V, KONOLIGE K, *et al*. ORB: an efficient alternative to SIFT or SURF[C]//2011 International Conference on Computer Vision. Barcelona: IEEE, 2011: 2564-2571

[19] BAY H, ESS A, TUYTELAARS T, *et al*. Speeded-up robust features (SURF)[J]. *Computer*

最新录用

*Vision and Image Understanding*, 2008, **110**(3): 346-359

[20] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: Binary Robust invariant scalable keypoints[C]//2011 International Conference on Computer Vision. Barcelona: IEEE, 2011: 2548-2555

[21] WANG R, BAO L L, CAI Y X. Image feature extraction and matching of augmented solar images in space weather[J]. *Chinese Journal of Space Science*, 2023, **43**(5): 840-852

[22] GUO Pengpeng. Research on Similarity Detection of Environmental Supervision and Governance Images Based on Siamese Neural Networks[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2023 (郭彭鹏. 基于孪生神经网络对环境监管治理图像的相似度检测研究[D]. 南京: 南京邮电大学, 2023)

[23] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, **60**(6): 84-90

[24] HE K M, ZHANG X Y, REN S Q, *et al*. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778

[25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations. San Diego: OpenReview.net, 2015

[26] 付小娜. 基于深度学习的太阳黑子群分类方法研究[D]. 昆明: 昆明理工大学, 2019

[27] 樊石鸣. 基于深度学习的太阳黑子群磁类型分类方法研究[D]. 昆明: 昆明理工大学, 2023

[28] HE Xinran. Research on Solar Flare Forecast Model Based on Deep Learning[D]. Beijing: University of Chinese Academy of Sciences (National Space Science Center), 2021 (何欣燃. 基于深度学习的太阳耀斑预报模型研究[D]. 北京: 中国科学院大学(中国科学院国家空间科学中心), 2021)

[29] ZENG Xunwen. Research on Magnetic Type Recognition of Sunspot Group and Solar Flare Prediction Based on Deep Learning[D]. Nanchang: Nanchang University, 2022 (曾洵文. 基于深度学习的太阳黑子群磁类型识别与太阳耀斑爆发预报研究[D]. 南昌: 南昌大学, 2022)

[30] YUAN Hanqing. Research on Detection Method of CME Based on Deep Learning[D]. Nanjing: Nanjing University, 2021 (袁汉青. 基于深度学习的日冕物质抛射物检测方法研究[D]. 南京: 南京大学, 2021)

[31] 冼祥贵. 基于深度学习的日冕物质抛射检测模型[D]. 昆明: 昆明理工大学, 2021

[32] GUO Shaocong. Research on Image Feature Extraction Method Based on CNN and Transformer[D]. Baoding: Hebei University, 2024 (郭少聪. 基于 CNN 和 Transformer 的图像特征提取方法研究[D]. 保定: 河北大学, 2024)

[33] LI Rong, WU Yingzhi, TIAN Qihui, *et al*. Solar flare forecasting model based on multi-modal feature fusion[J]. *Scientia Sinica Physica*, *Mechanica & Astronomica*, 2024, **54**(12): 129611 (李蓉, 吴颖智, 田奇辉, 等. 基于多模态特征融合的太阳耀斑预报模型[J]. 中国科学:

最新录用

物理学 力学 天文学, 2024, **54**(12): 129611)

[34] SONG Mingzhao. Research on Solar Filament Detection Method Integrating CNN and Transformer[D]. Kunming: Kunming University of Science and Technology, 2024 (宋明昭. 融合 CNN 与 Transformer 的太阳暗条检测方法研究[D]. 昆明: 昆明理工大学, 2024)

[35] YANG Peng. Research on Small Object Segmentation Algorithm Based on Hybrid CNN and Transformer[D]. Kunming: Yunnan Normal University, 2024 (杨鹏. 基于混合 CNN 和 Transformer 的小目标分割算法研究[D]. 昆明: 云南师范大学, 2024)

[36] DOU Fengping, XU Long, ZHAO Dong. Super-resolution of solar magnetograms based on uncertainty loss and attention mechanism[J]. *Acta Astronomica Sinica*, 2024, **65**(4): 63-78 (窦凤平, 徐龙, 赵栋. 基于注意力机制和不确定性损失的太阳磁图超分[J]. 天文学报, 2024, **65**(4): 63-78)

[37] MKWANDA N M, TIAN W X, LI J L. Sunspot group detection and classification by dual stream convolutional neural network method[J]. *Research in Astronomy and Astrophysics*, 2024, **24**(9): 095022

[38] GOPALSWAMY N, YASHIRO S, MICHALEK G, *et al*. The SOHO/LASCO CME catalog[J]. *Earth*, *Moon*, *and Planets*, 2009, **104**(1/2/3/4): 295-313

[39] CHEN T, KORNBLITH S, NOROUZI M, *et al*. A simple framework for contrastive learning of visual representations[C]//Proceedings of the 37th International Conference on Machine Learning. Vienna: PMLR, 2020: 1597-1607

[40] TÖRNROS M, BERRIOS D, BOCK A, *et al*. Interactive visualization of space weather data[C]//American Geophysical Union Fall Meeting. San Francisco: American Geophysical Union, 2012

最新录用

第一作者：

郭媛 女，2000 年 10 月出生于山西省晋中市，现为中国科学院国家空间科学中心学生，主要研究方向为空间环境应用。



通信作者：

包黎莉 女，1986 年 9 月出生于浙江省湖州市，现为中国科学院国家空间科学中心工程师，主要研究空间环境预报可视化、空间环境动态模拟仿真可视化。



最新录用