非协作卫星时空图卷积的微推力检测方法

王俊杰[1]，孙博文[2]，潘晓刚[2]，彭 永[1]，王炯琦[1]
（1. 国防科技大学 理学院，长沙 410000；2. 国防科技大学 系统工程学院，长沙 410000）

摘 要：针对非协作卫星存在的微弱推力因量级小、持续时间长，极易被观测噪声淹没，导致检测极为困难，在低轨卫星精密定轨与深空探测器轨道控制中途修正、近月制动等关键环节均普遍存在的问题，提出基于时序和空间学习的时空图卷积网络（Spatio-Temporal Graph Convolutional Network，ST-GCN）模型。通过构建 ST-GCN 模型，引入卷积网络（Convolutional Neural Network，CNN）、图神经网络（Graph Neural Network，GNN）、特征融合，并采用自适应学习率调度与图邻接权重的网络参数策略，以提高检测精度，有效解决了因卫星微推力极小导致难以检测的问题。通过3组低轨卫星微推力仿真数据集试验表明：与当前流行的时序模型相比，在综合精度和效率两个维度方面，本文提出的ST-GCN模型具有最优的综合性能。该方法具备向深空探测场景迁移的潜力，可为空间非合作目标态势感知提供技术参考。

关键词：微推力检测；时空图卷积神经网络（ST-GCN）；非协作卫星；时空建模；特征融合

中图分类号：V43 **文献标识码**：A **文章编号**：2096-9287(2026)04-0335-11

DOI:10.3724/j.issn.2096-9287.2026.20260028

引用格式：王俊杰，孙博文，潘晓刚，等. 非协作卫星时空图卷积的微推力检测方法[J]. 深空探测学报（中英文），2026，13（4）：335-345.

Reference format: Wang J J，Sun B W，Pan X G，et al. Non-cooperative micro-thrust detection method for satellites based on spatiotemporal graph convolution[J]. Journal of Deep Space Exploration，2026，13（4）：335-345.

引 言

卫星在轨道精密保持与机动控制中，微推力是太空抵近、轨道保持的主要手段[1-3]。然而，当卫星加速度降至微米级加速度水平时，微弱推力信号极易被观测噪声掩盖；持续数天甚至数月的推力会使轨道预报误差逐渐累积，最终引发碰撞风险或任务失败[4-8]。对于深空探测任务而言，探测器在中途轨道修正、近月制动、小行星绕飞等关键阶段同样面临类似的微弱推力检测与机动状态感知需求，其精度直接关系到任务成败[9]。因此，提升非合作目标微推力检测能力已成为空间态势感知与深空探测共同亟需突破的核心环节。

微推力难以检测的根源在于信号与噪声几乎处于同一量级：微米级加速度对应的单圈次漂移仅数十米，在数千米的轨道误差面前微不足道；恒定推力信号的能量被平摊至整个弧段，其频谱与大气阻力、太阳光压等保守力高度重叠；测距、测角噪声及时钟漂移进一步将信噪比压低至0 dB以下[10]。当上述特征叠加时，使得即便采用高精度轨道积分和扩展卡尔曼滤波[11]（Extended Kalman Filter，EKF），也难以在残差序列中形成可识别的系统性偏移，最终造成“推力存在却检测不到”的困境。传统轨道改进与滤波方法难以分离微推力弱信号，而机器学习技术可从海量数据中挖掘出关键特征，自主把握数据的潜在模式，并能针对新数据开展分析与归类。该技术为轨道异常识别带来突破，可高效精准检测异常，提升系统可靠性[12]。

以往研究主要分为以下两类。其一是基于物理模型主导的轨道机动检测方法。由美国战略司令部在Space-Track.org公开的两行根数（Two Line Elements，TLE）空间物体目录，该数据能够反映卫星的历史运动情况[13]。文献[14]提出了以 TLE 两行根数为基础的卫星机动检测方法，将轨道的半长轴作为特征参数之一，从而实现对卫星面内的机动检测。文献[15]利用SGP4模型长期项对目标特征轨道参数的预报值，与TLE真值之差建立阈值门限，检测地球低轨道（Low Earth Orbit，LEO）LEO的轨道机动事件。文献[16]在测元特性分析的基础上，对测距、测速等进行观测误

收稿日期：2026-04-21 修回日期：2026-06-15

基金项目：国家自然科学基金（62503487）；国防科技大学青年自主创新科学基金（XX25-60）；国防科技大学自主创新基金（24-XXXX-XXX-01-01）；空间智能控制技术全国重点实验室基金（2024-CXPT-GF-JJ-012-16）；航天器热控全国重点实验室开放基金（NKLST-JJ-005）

差建模。文献[17]基于小波分析理论，提出了固定分组法和移动开窗法两种方案，均能快速且准确地检测出地球同步轨道（GEostationary Orbit，GEO）卫星异常。目前物理模型主导的检测方法精度和可解释的优势仍在，但当加速度降至微米级、单圈次漂移约数10 m时，其模型精度、观测采样、特征维度和检测策略均不足以从TLE噪声中提取弱信号。

第二类则是基于数据驱动主导的轨道机动检测方法，文献[18]针对GEO卫星未来机动状态识别问题，提出了一种“经度预测+机动识别”的自适应深度学习方法。该方法把经度预测最小平均误差、均方根误差平均降低一个数量级。文献[19]采用非监督分类方法的K-均值、层次、模糊 C-均值聚类处理两行根数的历史数据并识别轨道机动。文献[20]这项工作提出了一种基于优势执行评价器（Advantage Actor-Critic，A2C）算法的强化学习框架，用于优化卫星轨道参数，以在预定义的表面半径内实现精确的地面覆盖。文献[21]利用长期短期记忆（Long Short-Term Memory，LSTM）神经网络，通过操作对象的生命模式（包括其状态和机动历史），对新传入的轨道参数序列进行机动检测。目前，基于机器学习的卫星轨道异常识别研究比较火热，涌现出大量算法。

本文基于此，构建了时空图卷积网络模型（Spatio-Temporal Graph Convolutional Network，ST-GCN）。该模型通过链式图将时间步转化为节点，使图卷积天然具备时序传播能力，实现了对时序和空间的联合学习，填补了现有方法均未覆盖的空白。该模型的首次提出主要应用于交通预测和人体骨架识别方面，本文把该模型迁移应用于轨道微推力识别。该模型包括3个主要模块：数据处理、时空图学习和分类预测。在数据处理模块中构建时序无向图结构，将相邻时间步作为图的边，把一维时序数据转化为图数据，适配卫星数据的时空依赖性。用一维卷积神经网络（1D Convolutional Neural Network，1D-CNN）和GCN 堆叠的结构来简化时空图学习模块，以提高模型运行效率。训练用损失函数、预测准确率以及均方误差（Mean Square Error，MSE）和平均绝对误差（Mean Absolute Error，MAE）量化模型效果。

值得说明的是，本文的实验验证基于低轨卫星场景。这主要基于两方面考虑：其一，低轨卫星密集、地基雷达观测数据丰富，是验证非合作目标机动检测方法的理想试验场；其二，低轨卫星所面临的微弱推力检测问题，在数学本质和工程挑战上与深空探测器的中途轨道修正、近月制动等机动检测问题高度一致。本文所提的时空图卷积检测框架具有轨道高度无关性，可直接迁移至深空探测场景。此外，低轨卫星作为深空探测天基测控站和通信中继节点，其自身机动状态的准确感知对深空测控网的可靠运行同样具有重要意义。

1 问题定义

给定一条卫星过境弧段的多参数观测序列 $\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_T]\in R^{T\times 5}$，其中 $\boldsymbol{x}_t=[t,az_t,el_t,r_t,\rho_t]$ 分别表示Min-Max归一化后的时间戳、方位角、俯仰角、站星距离及站星距离变化率，$T=100$为统一弧长。地面探测器可观测区域示意图如图1所示。

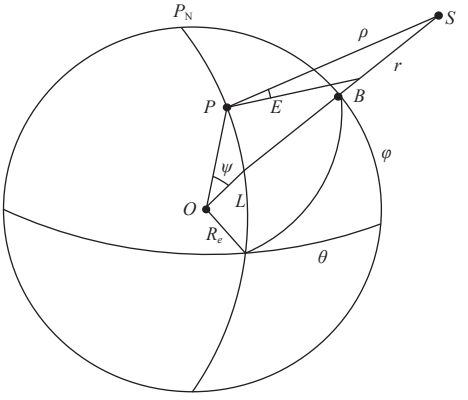


图 1 地面探测器可观测区域

Fig. 1 Observation area of ground detector

图1中点O为地心，点P为观测点，点S为卫星，点B为卫星的星下点，R_e为地球半径，r为卫星轨道半径，ρ为站星距离，θ为星下点与观测点的经度差，φ为星下点的纬度，ψ为观测站与星下点的地心角，L为观测站的纬度，E为观测仰角。

定义图$\mathcal{G}=(\mathcal{V},\mathcal{E})$，节点$\mathcal{V}$对应100个时间步长，时间步长设置为10 s，边集采用时间邻接$\boldsymbol{\varepsilon}=\{(i,i+1),(i+1,i)\mid i=1,2,\cdots,T-1\}$，即仅在相邻的epoch间建立双向边，以刻画轨道状态的马尔可夫特性。其中，标签$y\in\{0,1,2,3\}$分别表示“无推力”“x向推力”“y向推力”“z向推力”时的状态。目标为学习映射f：$X\rightarrow y$，实现对1×10^{-6}～9×10^{-6} m/s^2、持续 100 s的微米级推力单圈次识别。

2 时空图卷积网络模型

2.1 总体框架设计

时空图卷积网络模型如图2所示。模型输入为卫星过境弧段Min-Max归一化后的观测序列，即时间、方位角、俯仰角、站星距离、站星距离变化率，输出为模型的预测类型。核心目标是实现非协作低轨卫星微推力的存在性检测与方向识别。

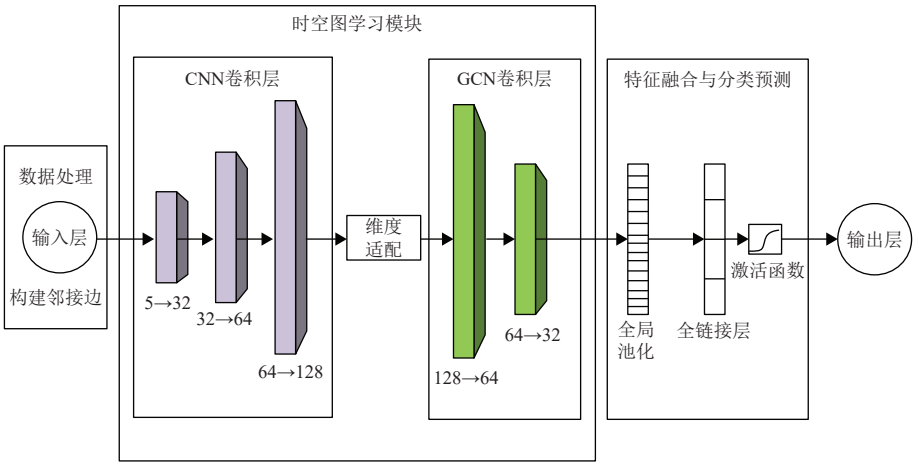


图 2　ST-GCN模型示意图

Fig. 2　Schematic diagram of ST-GCN model

时空图学习模块包含时间和空间学习；时间学习模块通过1-D CNN提取多元时间序列的局部时序特征，经过维度适配后把时序特征转为图的节点特征，再通过空间学习模块通过GCN聚合相邻时间步的状态信息，链式连接后实现时空特征的联合建模。

2.2　输入与图构建

模型输入为Min-Max归一化后的5维观测序列$\boldsymbol{X}\in R^{\wedge}\{T\times 5\}$，其中$T=100$为统一时间步长，每个时间步对应10 s的观测间隔，5个特征分别为时间、方位角、俯仰角、站星距离及站星距离变化率。

将每个时间步视为图中的一个节点，节点特征为对应时刻经过CNN提取的高维抽象特征。该抽象特征融合了原始5个物理量的非线性组合、局部时间窗口的上下文信息、多层卷积提取的不同层次的模式。

边连接采用一阶时间邻接策略，即在相邻时间步之间建立双向边，形成链式图结构，使信息在这些节点之间进行传递，进一步建模时间序列的全局依赖关系。在训练过程中，邻接矩阵的边权重固定为1。与全连接拓扑图相比，链式图参数减少（T-1）×（T-2）/2个边，其中T为节点数，本文T取100。除此之外链式图还能避免过拟合，更适用于长时序数据的处理。

2.3　时序分支

传统的CNN只能在局部感受野内工作，因此对于较长的序列建模和处理长期依赖关系表现较差。为此，引入了1-D CNN来扩大感受野且不增加计算量，通过扩张因子（dilation rate）使得卷积核每次滑动时，跳过若干个位置，使得每个卷积核能覆盖更多的时间步，达到扩大感受野的目的[22]。

设F：$Z^2\to R$为离散函数。设$\Omega_r=[-r,r]^2\cap Z$，k：$\Omega_r\to R$是一个离散滤波器，其大小为$(2r+1)^2$。离散卷积算子可以定义为

$$(F*k)(p)=\sum_{s+t=p}F(s)k(t) \tag{1}$$

其中：F为时序特征离散函数；k为卷积核；p为特征位置。现在推广这个算子，设l是扩张因子，通过间隔采样扩大感受野，设$*_l$定义为

$$(F*_l k)(p)=\sum_{s+l*t=p}F(s)k(t) \tag{2}$$

把$*_l$称为扩张卷积或l-D 扩张卷积。当扩张因子$l=1$时，扩张卷积*退化为普通的离散卷积[23]。为捕获微米级推力在时序的微弱变化，采用三层一维卷积网络，每层卷积核大小为3，填充为1，以保证时间维度的长度不变。卷积通道数逐层递增，由32到64再到128，以逐步提取更高层次的时序特征。每层卷积后接BatchNorm、ReLU激活函数和Dropout，以防止过拟合并加速训练。3层膨胀卷积（dilation）的膨胀因子分别设为1、2、4，单层卷积核大小为3，总感受野长度为$1+(3-1)\times(1+2+4)=15$个时间步，感受野扩大至150 s，可覆盖100 s的微推力持续窗口。其结构如图3所示。

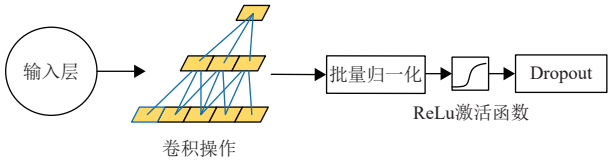


图 3　卷积层示意图

Fig. 3　Convolution layer diagram

2.4　空间分支

GCN为专为处理图结构数据而设计的一种神经网

络架构，能融合图拓扑关系同节点的特征，并经图卷积运算从而挖掘节点间关系与特征。该算法的关键点是基于邻域信息的聚合，即每个节点不仅受自身特征的影响，还会通过与相邻节点信息的交互来更新。这一信息传播机制使节点可整合邻居信息，进而捕捉时序数据中的先后顺序。其结构如图4所示。

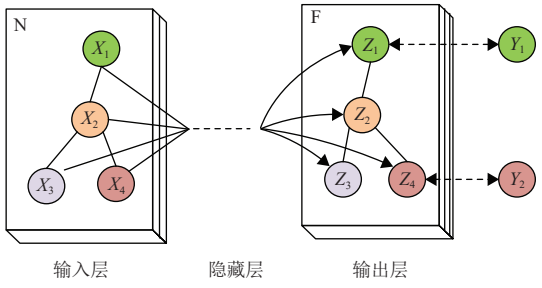


图 4 图卷积神经网络[24]

Fig. 4 Graph convolutional neural network[24]

由图4可知，图结构数据的输入层包含N个节点，经中间隐藏层处理，最终输出层生成F个节点特征。输入层用于提取原始特征，图卷积网络则是用于多层级特征的聚合，最后输出层输出分类结果。GCN基于谱图理论定义了图卷积操作，并通过近似简化使得其高效地在空间域中进行计算。图卷积的核心公式为

$$\boldsymbol{X}^{(l+1)} = ReLU\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{X}^{(l)}\boldsymbol{W}^{(l)}\right) \tag{3}$$

其中：$\boldsymbol{X}^{(l)}$为第(l)层的节点特征矩阵，输入层（$\boldsymbol{X}^{(0)}=\boldsymbol{X}$），即原始特征矩阵；$\tilde{\boldsymbol{A}}$为加了自环的邻接矩阵，即$\tilde{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{I}$，其中$\boldsymbol{I}$为单位矩阵；$\tilde{\boldsymbol{D}}$为$\tilde{\boldsymbol{A}}$的度矩阵，$\tilde{\boldsymbol{D}}_{ii}=\sum_{k=1}^{k=N}\tilde{\boldsymbol{A}}_{ik}$，非对角元为 0；$\boldsymbol{W}^{(l)}$为第$(l)$层的可学习权重矩阵；$ReLU(\cdot)$为ReLU激活函数。

$\tilde{\boldsymbol{A}}$为节点加上自环，使节点能使用自己的信息。$\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}$实现了对邻接矩阵的标准化，从而防止度数不同的节点对结果产生较大的影响。$H^{(l)}W^{(l)}$表示线性变换，提取特征。$ReLU(\cdot)$为模型增加非线性变换，以提升模型的表达能力。

该模型中，通过两层GCN的堆叠，节点能够与更多的距离更远的邻居产生融合，获取特征信息，扩大感受视野，从而能识别出多种特征之间长距离的时空关系。

2.5 特征聚合与分类

该阶段将前面CNN提取的时间特征和 GCN提取的空间特征进行最终的全局聚合与高层语义融合，把低层次的时空特征转化为可用于决策的分类概率，是连接特征提取与任务输出的关键环节。

首先，该阶段将GCN输出的特征沿时间维度进行全局均值池化。通过把任意长度的卫星轨道观测弧段压缩为固定长度的嵌入向量，既保留了全局上下文的多维度特征信息，又消除了数据缺失带来的尺度变化。

设GCN输出的节点嵌入矩阵为$\boldsymbol{H}\in\mathbb{R}^{T\times 32}$。为获得全局的序列表示，沿时间维度进行全局均值池化，即

$$\boldsymbol{e}=\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{h}_t\in\mathbb{R}^{32} \tag{4}$$

其中：$\boldsymbol{h}_t$为$\boldsymbol{H}$的第t行，对应时间步t的嵌入向量。嵌入向量$\boldsymbol{e}$先经过Dropout层以防止过拟合，随后输入全连接层。

$$\boldsymbol{z}=\boldsymbol{W}_{\mathrm{fc}}\boldsymbol{e}+\boldsymbol{b}_{\mathrm{fc}}\in\mathbb{R}^{4} \tag{5}$$

其中：$\boldsymbol{W}_{\mathrm{fc}}\in\mathbb{R}^{4\times 32}$，$\boldsymbol{b}_{\mathrm{fc}}\in\mathbb{R}^{4}$为可学习参数。最后采用log-softmax函数输出4个分类的对数概率。

$$\hat{\boldsymbol{y}}=\log\left(\frac{\exp(\boldsymbol{z})}{\sum_{c=1}^{4}\exp(\boldsymbol{z}_c)}\right) \tag{6}$$

训练阶段使用负对数似然损失（NLLLoss）作为损失函数来保证分类正确率。

$$\mathcal{L}=-\frac{1}{N}\sum_{i=1}^{N}\log\hat{y}_{i,y_i} \tag{7}$$

其中：N为batch大小，batch size = 256；y_i为第i个样本的真实标签。推理阶段仅需一次前向传播即可得到分类结果，计算开销可忽略不计。

2.6 训练策略

2.6.1 优化器的选择

在深度学习中，优化器是训练过程中的重要组成部分。不同的优化器可通过不同的策略来调整学习率和更新规则来帮助模型收敛得更快、更好。该模型使用AdamW优化器（adaptive moment estimation with weight decay），其核心公式为

$$m_t=\beta_1 m_{t-1}+(1-\beta_1)g_t \tag{8}$$

$$v_t=\beta_2 v_{t-1}+(1-\beta_2)g_t^2 \tag{9}$$

$$\hat{m}_t=\frac{m_t}{1-\beta_1^t},\quad \hat{v}_t=\frac{v_t}{1-\beta_2^t} \tag{10}$$

$$\theta_t=\theta_{t-1}-\frac{\eta_t}{\sqrt{\hat{v}_t}+\epsilon}\cdot\hat{m}_t-\eta_t\lambda\theta_{t-1} \tag{11}$$

其中：m_t为梯度的一阶矩即均值；v_t为梯度的二阶矩即方差；$\hat{m}_t$、$\hat{v}_t$为对m_t、v_t进行偏差修正后的估计；g_t为当前时间步的梯度，$g_t=\nabla_\theta L(\theta_t)$；$\beta_1$和$\beta_2$为一阶矩和二阶矩的衰减率；$\epsilon$为一个小常数，用于避免除零错误；$\eta_t$为当前时刻的学习率；$\lambda$为权重衰减系数。

AdamW优化器的核心思想在于把梯度更新与权重衰减分离开来，改为显式更新，使得权重衰减能够正确地按层进行处理。即它将权重衰减视为一种对权重的直接约束，而不是将其包含在梯度计算中。通过将梯度更新与权重衰减分离开来，AdamW优化器可以更加有效地控制权重的更新步长，以避免在训练过程中出现权重过大或过小的情况，从而提高了模型的收敛速度和稳定性。除此之外，权重衰减还是一种正则化方法，可以有效地防止模型过拟合。

2.6.2　学习率调度

一般来说，在模型训练过程中，通常期望在训练初期设置较高的学习率，以便网络可快速收敛；而在训练后期，则将学习率降低，帮助网络能够更精准地接近最优解。初始学习率为0.000 5，衰减为先慢中快后慢，以提高计算效果。本文使用的是余弦退火策略，以平滑收敛为目的，即学习率随训练迭代次数呈余弦式衰减的方法，学习衰减曲线如图5所示。

$$\eta_t=\eta_{\min}+\frac{1}{2}(\eta_0-\eta_{\min})\left[1+\cos\left(\frac{t\pi}{T_{\max}}\right)\right] \tag{12}$$

其中：η_t为第 t 次迭代时的学习率；η_0为初始学习率；$\eta_{\min}$为最小学习率，$\eta_{\min}=0.000\,1$；$T_{\max}$为最大迭代次数，$T_{\max}=300$。

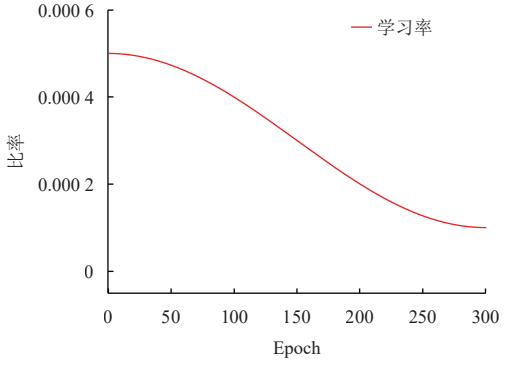


图 5　学习率衰减曲线

Fig. 5　Learning rate decay curve

2.6.3　早停机制

早停法（early stopping）是一种用于避免过拟合的训练策略。当验证集上的loss不再显著下降或变化幅度低于预设阈值时，训练就会提前终止，从而提高模型效率[25]。

停止标准：设E为训练模型的误差函数，$Etr(t)$是训练数据的误差，那么$Eval(t)$是验证集的误差。其实在实际情况下往往无法直接知晓模型的泛化误差具体数值。因此，通常只能通过验证集的误差来对其进行近似估计。假设$Eopt(t)$是迭代次数t时，取得的最好验证集误差为

$$Eopt(t)=\min_{t'\leqslant t}Eval(t') \tag{13}$$

引入一个新的指标，命名为泛化误差（generalization loss），用来表示在当前迭代周期下，当前的泛化误差与迄今为止记录的最低误差之间的相对增长比例。

$$GL(t)=100\times\left(\frac{Eval(t)}{E_{\text{opt}}(t)}-1\right) \tag{14}$$

若泛化损失值偏高，往往意味着模型出现了过拟合的情况。一旦泛化损失超出设定的阈值范围，则终止训练过程。在此用GL_α来定义，当GL_α大于预设阈值α_{th}时，停止训练，这里$\alpha_{\text{th}}=10$。为避免训练初期验证损失波动导致的误判，设置前50个epoch为预热期，预热期内不执行早停，若$GL_\alpha\geqslant 10$且连续10个epoch均未改善，则终止训练。

3　仿真实验与模型效能评估

首先构建数据集，数据集为仿真生成的地基雷达观测得到；然后分别对比验证ST-GCN 和其它模型识别效果；最后进行消融实验对方法进行测试。

3.1　数据集构建

每个数据样本共包含5类观测数据，分别为时间、方位角、俯仰角、站星距离、站星距离变化率，仿真时长为1 000 s，步长为10 s，共生成100个时间步的观测序列。每个数据集中包含 3 200个数据样本。

方位角是从某点的指北方向线起依顺时针方向至目标方向线间的水平夹角，用“度”表示。在部分数据集中的任意时刻随机加入数值在$1\times10^{-6}-9\times10^{-6}$ m/s^{2}持续时间为100 s的x、y、z这3个方向的微推力。坐标系的原点为卫星质心，沿轨道径向指向地心的方向为x轴，沿轨道切向的方向为y轴，z轴垂直于轨道平面，构成右手坐标系。

所构建的每个数据名称会标注是数据类型，且所有数据集中均添加了噪声作为干扰。数据的名称为所添加微推力动力值的大小，即0123：1×10^{-6}～3×10^{-6} m/s^2；0345：3×10^{-6}～5×10^{-6} m/s^2；0369：3×10^{-6}～9×10^{-6} m/s^2。

该数据集通过仿真代码进行轨道模拟生成，神经网络则使用 Python 3.8 和 PyTorch 2.4.1深度学习框架

实现。模型训练在一台运行Windows 11的电脑进行，该电脑配备了支持CUDA 12.8加速的NVIDIA RTX 4060 GPU，以及32 GB内存。

受真实的地基雷达观测非协作卫星数据难以获取的限制，本文仅采用高保真仿真数据验证方法有效性。未来将对模型进行调整后，利用STARLINK公布的TLE数据开展实测验证。

3.2 仿真效验

为更直观地对模型的效能以及各方面有个充分的认识，在此引用较为常见的TCN、LSTM、GRU、Transformer4种主流时序模型作为基线并输入相同的数据集进行对比。

3.2.1 同一数据集下的模型检测

5种算法的网络信息参数记录在表1中。根据表1可知，ST-GCN模型的体量与GRU模型相当，小于其它模型，说明预测卫星机动时直接套用以往的网络会造成不必要的计算资源浪费，证明了本文简化的ST-GCN模型的简洁性。为更好地进行对比，5类算法全部使用相同的损失函数、优化算法、初始化学习率、正则化参数。

表 1 网络模型信息

Table 1 Network model information

特征参数	ST-GCN	TCN	LSTM	GRU	Transformer
可学习参数总数	42340	164356	53296	40560	76068
损失函数	NLL	NLL	NLL	NLL	NLL
初始学习率/%	0.0005	0.0005	0.0005	0.0005	0.0005
丢弃率/%	0.3	0.3	0.3	0.3	0.3
时间步长	100	100	100	100	100

3.2.2 结果对比

本节对ST-GCN模型进行评估。根据3.1小节生成的LEO卫星运动规律数据集，分别使用两种模型进行测试。5种方法针对数据集的平均效果见表2。使用Accuracy准确率、MSE均方差、MAE平均绝对误差作为评价指标。计算公式为

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \tag{15}$$

其中：TP为实际是正样本，模型预测为正样本的数量；TN为实际是负样本，模型预测为负样本的数量；FP为实际是负样本，模型预测为正样本的数量，即误判数量；FN实际是正样本，模型预测为负样本的数量，即漏判的数量。

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2 \tag{16}$$

$$MAE = \frac{1}{m}|y_i - \hat{y}_i| \tag{17}$$

其中：y_i为真实微推力特征值；$\tilde{y}_i$为模型预测的微推力特征值。

表 2 多种方法的平均性能

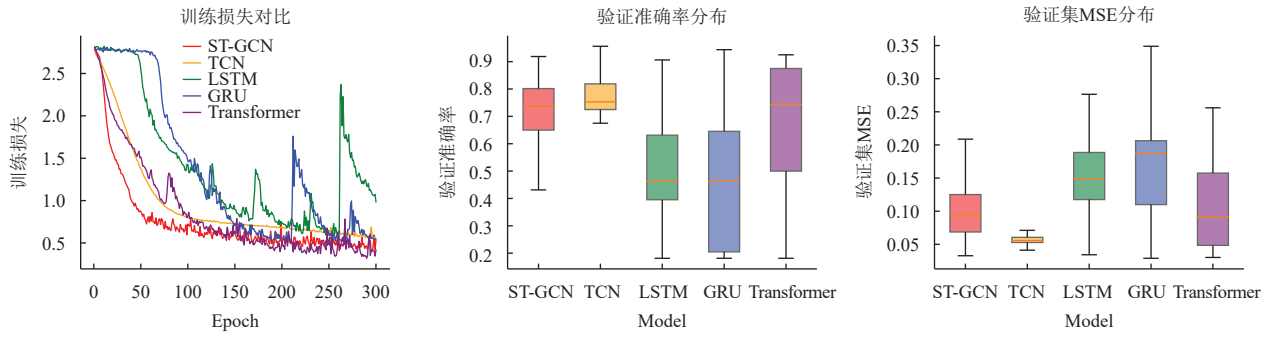
Table 2 Average performance of two methods

数据集	指标	ST-GCN	TCN	LSTM	GRU	Transformer
0123	MSE	0.032	0.057	0.0413	0.0373	0.0356
	MAE	0.0636	0.1353	0.0927	0.0803	0.0777
	ACC/%	93.5	86.2	87.12	89.38	91.5
	时间/s	57.9	66.9	56.05	103.57	74.37
0345	MSE	0.0249	0.021	0.0370	0.0302	0.0213
	MAE	0.0561	0.0512	0.0782	0.0577	0.0463
	ACC/%	95	91.75	88.12	92.5	93.75
	时间/s	57.2	65.13	55.86	111.65	79.97
0369	MSE	0.0272	0.037	0.0388	0.0281	0.0196
	MAE	0.0712	0.0808	0.0851	0.0546	0.0410
	ACC/%	94.75	89.3	88.75	93.13	95
	时间/s	57.2	68.5	55.47	98.87	74.21

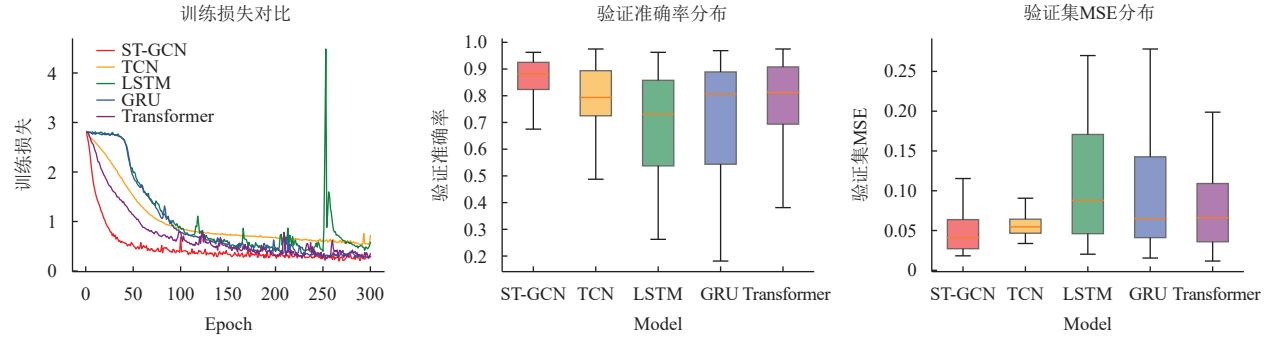
时间是指训练300个epoch后的模型运行一次需要消耗的计算时间，单位为秒（s）。输入数据集0123、0345、0369后的训练过程如图6所示。根据不同数据集的训练效果图发现，ST-GCN模型更稳定、损失函数收敛更平稳、预测效果更准确、预测误差更小。根据表2可知，在准确率方面，ST-GCN在3个数据集中ACC分别为93.5%、95%、94.75%，均优于TCN、LSTM和GRU。与Transformer相比，ST-GCN在0123和0345数据集上分别高出2%和1.25%，在0369数据集略低0.25%。平均而言，ST-GCN的ACC值达94.41%，高于其它时序模型，表明时空联合建模对微弱推力信号的判别能力优于纯时序模型。

在回归误差方面，ST-GCN在0123和0369数据集的MSE/MAE均优于TCN和LSTM。但在0345数据集，ST-GCN的MSE和MAE略高于TCN。主要原因为ST-GCN通过图卷积对相邻时间步的特征进行平滑聚合，增强了模型对噪声的鲁棒性，从而提升了推力存在性的判别能力；但平滑操作同时削弱了对微弱推力幅度变化的敏感性，导致回归误差略大。而TCN的因果卷积对局部突变更敏感，回归占优但分类效果稍差。鉴于本任务的核心是推力识别，ST-GCN在ACC的优势更具实际意义。由于在数据集0345中ST-GCN模型的MAE和MSE值均高于TCN模型，现进行显著性检验。用已经训练好的模型从数据集中随机抽取800个样本进行检测，其误差分布如图7所示。平均而言ST-GCN的误差是小于TCN的，但存在少数离群值较大的情况。经计算后p值为$8.106\,6\times10^{2}$，无法拒绝原假设，即两个模型的误差分布无显著差异（$p >= 0.05$）。

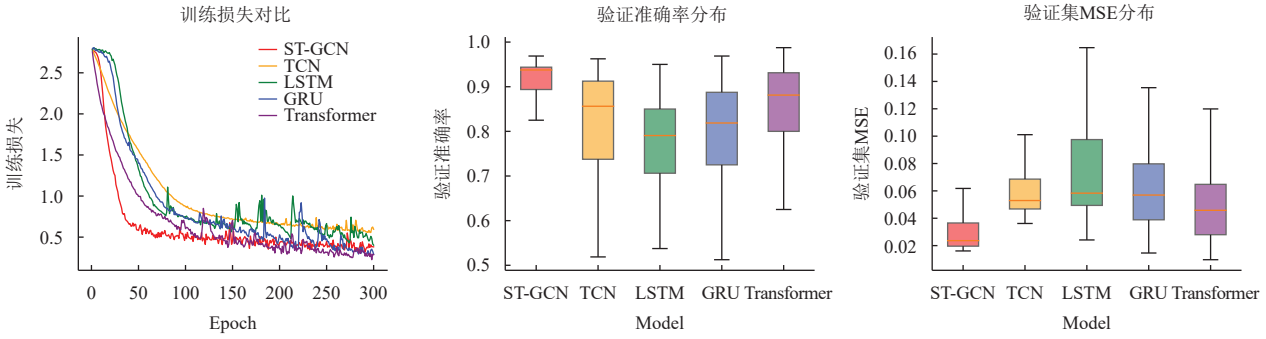
在推理时间方面，ST-GCN 模型展现出优异的实时性能，3个数据集平均推理时间仅为 57.4 s，仅比最



（a）数据集0123下模型效果对比



（b）数据集0345下模型效果对比



（c）数据集0369下模型效果对比

图 6　不同数据集下的运行效果

Fig. 6　Operating performance on different datasets

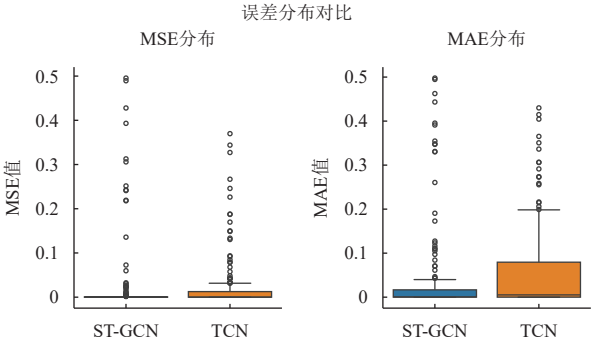


图 7　消融模型效果对比

Fig. 7　Performance comparison in ablation study

快的LSTM模型慢3.2%，但比 Transformer快24.4%，比GRU快45%。这一结果表明，ST-GCN 在保持高精度的同时，能够实现高效计算。

综合精度和效率两个维度，本文提出的ST-GCN模型具有最优的综合性能。ST-GCN在分类准确率方面全面领先，回归精度接近最优的 Transformer，但计算速度比Transformer快近1/4，适合对实时性要求高的卫星地面监测系统。ST-GCN在3个不同特性的数据集性能波动最小，没有出现TCN那样的极端表现，能适应不同机动类型的卫星数据。

3.3　消融实验

为验证面向LEO卫星微推力识别的ST-GCN方法关键策略的有效性，本文将系统地开展消融实验。在检测方法中，以方位角、俯仰角、站星距离及站星距离变化率4维时序特征为主体输入；随后构建ST-GCN框架进行同一弧段内的异常概率预测，该框架通过链式图卷积聚合相邻时间步状态，并引入1-D CNN捕获长

程时序依赖。

先依次剥离核心组件：去除图卷积模块、去除时序CNN模块、移除残差连接、替换链式图为全连接拓扑图，分别评估对检出率与虚警率的影响；进一步对时间窗口长度、邻接阈值及Dropout率等超参数进行遍历，以确定最优配置。最后，结合平滑误差与自适应阈值更新策略，量化验证各组件对模型的边际贡献，证明所提时空联合建模及链式图结构在提升LEO卫星微弱推力检测稳健性方面的必要性与有效性。

3.3.1 模型结构修改

为验证时空图学习模块中的时间学习模块CNN与空间学习模块的GCN的必要性，本小节将会构建4个模型，分别为A1、A2、A3、A4。A1为原模型，A2为移除空间学习模块后的模型，A3为移除时间学习模块后的模型，A4为移除残差连接的模型。本节仍以ACC、MSE、MAE、时间作为评价指标，并统一使用数据集0345作为输入，得到的结果如表3所示。

表 3 消融模型效果对比表

Table 3 Ablation experiment results

模型	MSE	MAE	ACC/%	时间/s
A1	0.024 9	0.056 1	95.00	57.20
A2	0.041 5	0.108 9	93.25	41.95
A3	0.189 5	0.374 5	26.87	37.98
A4	0.540 0	0.094 6	92.50	57.70

由表3和图8可知，A2模型的准确率有较小幅度的下降，MSE和MAE都增大了1倍，代表着模型拟合效果下降。由于A2模型砍去了GCN部分，因模型简化，故总体运算时间减少。A3模型几乎不收敛，准确率大幅降低，表示单纯的GCN模型并不能很好地学习多元时间序列数据。A4模型的准确率有较小幅度的下降，MSE增大了近26倍，代表在模型运行中出现了异常值，导致MSE波动较大。推测是因网络梯度消失导致局部预测值偏离真实值，出现异常误差，验证了残差连接对模型稳定性的关键作用。

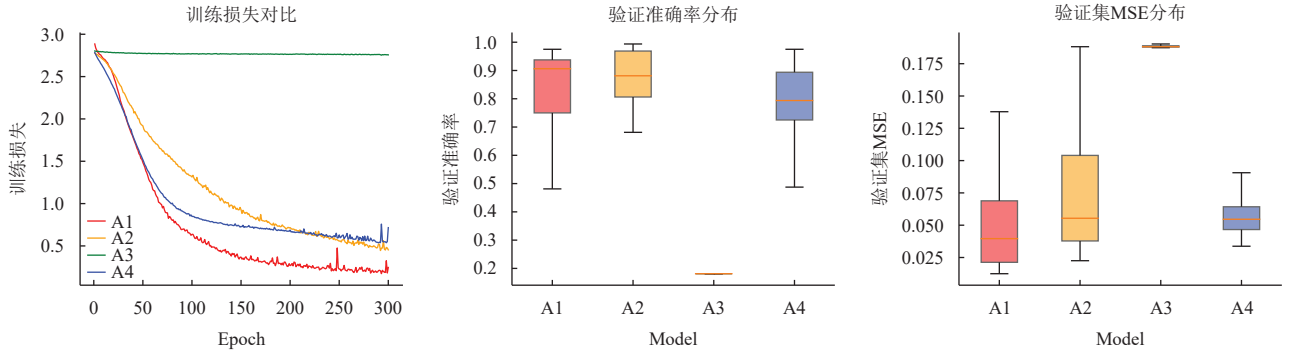


图 8 不同图连接方式的效果对比

Fig. 8 Comparison of effects of different graph connection methods

由图9所示，大部分的样本残差落在[0, 0.2]区间内，且围绕零残差线对称。这表明模型对大部分测试样本能够正确分类且置信度较高。但A4模型出现多个残差绝对值超过0.5、甚至接近–1的离群点，根据颜色可知主要为类别1的点，对应的是x轴机动样本。说明残差连接缺失后，模型对特定推力方向的检测能力出现了退化。

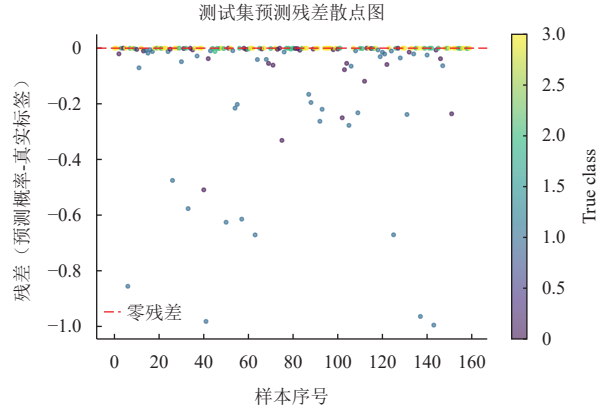


图 9 不同遗忘率的模型效果

Fig. 9 Effects of dropout rates on model performance

3.3.2 超参数组

（1）图连接方式

为探究图拓扑连接方式对机动检测的影响，分别构建一阶时序邻接、二阶邻接与全连接拓扑图开展消融实验。由图10可知，一阶邻接收敛速度更快，全连接虽训练集拟合最优但易过拟合；从箱线图和表4来看，一阶邻接准确率、MSE中位数最优且稳定性更强；二阶邻接依靠间接边小幅优化回归误差，却因冗余连接降低分类精度；全连接误差离散程度最大、泛

化性能较差。测试结果表明，一阶邻接准确率95%，高于二阶、全连接，推理耗时57.2 s；二阶邻接MSE、MAE略优但分类性能受损，全连接预测误差偏高且推理耗时高达278.32 s。综合检测精度、模型鲁棒性与时效性，一阶邻接是本任务最优的图构建方案。

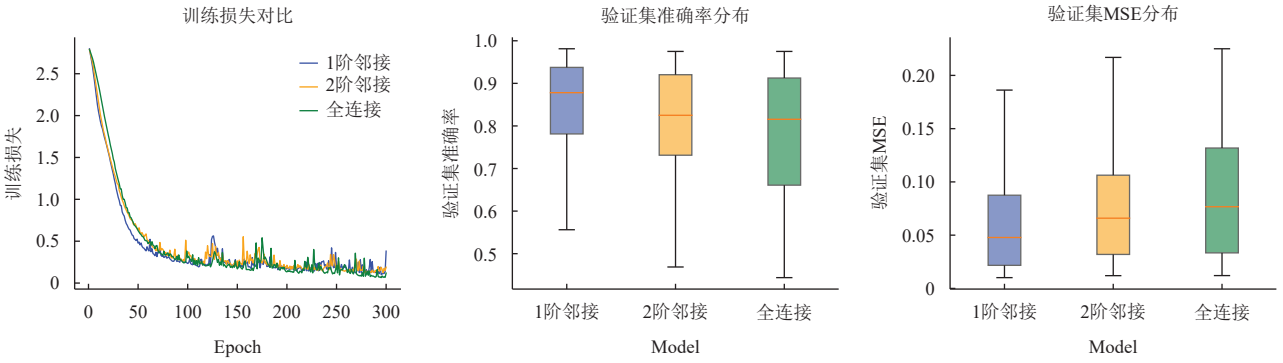


图 10　不同图连接方式的效果对比

Fig. 10　Comparison of effects of different graph connection methods

表 4　不同图连接方式

Table 4　Different graph connection methods

方式	MSE	MAE	ACC/%	时间/s
1阶邻接	0.024 9	0.056 1	95.00	57.20
2阶邻接	0.024 7	0.048 6	93.75	57.07
全连接	0.031 0	0.048 4	92.50	278.32

（2）Dropout

根据不同遗忘率下的效果表5和不同遗忘率的模型效果图11可知，当Dropout值为 0.3时，模型预测准确率最高，方差最小，抑制过拟合最佳。

表 5　不同遗忘率下的效果

Table 5　Model effects under different dropout rates

遗忘率	MSE	MAE	ACC/%	时间/s
0.3	0.024 9	0.056 1	95.00	57.2
0.5	0.067 0	0.105 5	95.10	57.2
0.7	0.078 8	0.120 5	93.13	55.7

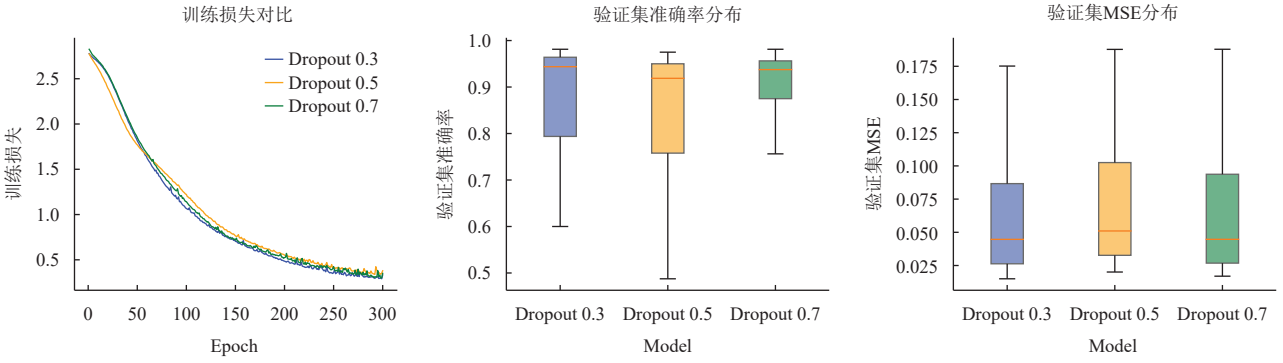


图 11　不同遗忘率对比图

Fig. 11　Comparison of effects under different dropout rates

为了探寻模型究竟关注哪些输入特征，现展开输入扰动分析。该实验旨在评估训练好的ST-GCN模型对输入数据扰动的鲁棒性。实验将原始测试集分别施加五种扰动：高斯噪声（标准差0.01、0.05、0.10）以及针对方位角、俯仰角、距离、距离变化率4个特征的随机掩码，掩码率为20%，即该维度有20%的数据会被归零。并与无扰动的干净测试集进行对比。所有数据均使用与训练集相同的标准化参数进行处理，确保公平性。

如图12所示，模型显示模型对微小随机噪声具有一定鲁棒性，但模型对较大幅度的输入波动较为敏感；方位角、俯仰角和站星距离变化率这几个特征对模型决策至关重要，一旦缺失会严重降低模型的检测效果。而距离特征相对次要，即使被随机掩码20%，模型仍能从角度和速度的时序模式中提取足够信息，维持近半数的正确分类能力。这一结果也为特征选择或传感器抗干扰设计提供了明确指导：应重点保障方位角、俯仰角和距离变化率的数据质量。

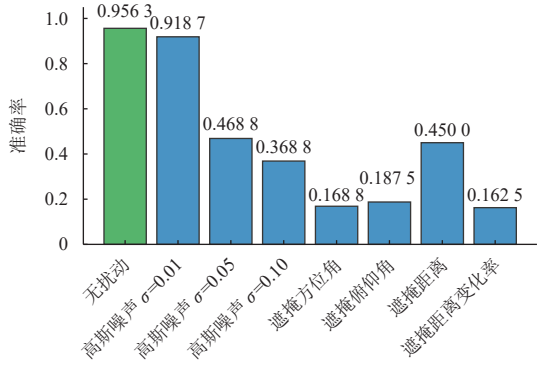


图 12 输入扰动下的模型鲁棒性

Fig. 12 Model robustness under input perturbations

4 结 论

本文将ST-GCN迁移至LEO轨道微推力检测领域，构建以时间步为节点、以邻接为顺序的链式图，兼顾轨道马尔可夫性与计算轻量化，设计膨胀CNN和残差GCN串行骨架。以方位角、俯仰角、站星距离、站星距离变化率四维测元为主体，构建时空互补信息；引入概率级MSE损失、MAE作为评估指标，显著提高识别准确率。在自建仿真数据集上开展消融与对比，所提方法对所构建的0123、0345、0369三组仿真数据集的机动检出率分别为93.5%、95.5%、94.75%。综合精度和效率两个维度，本文提出的ST-GCN模型具有最优的综合性能。

未来工作将重点围绕以下三个方向展开：首先，由于目前暂无地基雷达观测的真实数据，后续将调整模型输入结构，进一步应用于TLE或GNSS实测轨道数据，全面验证其在真实噪声环境下的鲁棒性与泛化能力；同时，针对GEO卫星的观测特性，完成对更长弧段及不规则采样间隔的算法适配；此外，将研究扩展至多目标联合检测场景，充分发挥图结构在建模复杂关系上的天然优势，通过刻画卫星间的相对运动关联，实现星群级的协同异常监测与态势感知。

参考文献

[1] Wang B C，Li S，Mu J Z，et al. Research advancements in key technologies for space-based situational awareness[J]. Space：Science & Technology，2022，2022：1-31.

[2] Liu D，Zong Q，Zhang X Y，et al. Enhancing space-based situational awareness：real-time observation of dynamic targets with MetaCooperative-Scheduling net[J]. IEEE Transactions on Aerospace and Electronic Systems，2024，60(6)：8198-8211.

[3] Liu A R，Xu X L，Xiong Y Q，et al. Maneuver strategies of starlink satellite based on space X-released ephemeris[J]. Advances in Space Research，2024，74(7)：3157-3169.

[4] Escribano G，Sanjurjo-Rivo M，Siminski J A，et al. A filter for tracking non-cooperative low-thrust satellites using surveillance radar data[J]. Acta Astronaut，2023，213：694-707.

[5] Zhang Z X，Shu L Z，Zhang K K，et al. Orbit determination and thrust estimation for noncooperative target using angle-only measurement[J]. Space：Science & Technology，2023(2023)：1-18.

[6] Singerman P G，O'Rourke S M，Narayanan R M，et al. Symbolic dynamics for radar target maneuver detection with high data rates[J]. IEEE Transactions on Aerospace and Electronic Systems，2024，60(2)：1647-1659.

[7] Zhang P，Wu D，Baoyin H X. Real-time hybrid method for maneuver detection and estimation of non-cooperative space targets[J]. Astrodynamics，2024(8)：437-453.

[8] Li T，Li K B，Chen L. Maneuver detection method based on probability distribution fitting of the prediction error[J]. Journal of Spacecraft Rockets，2019，56(4)：1114-1120.

[9] Yu S X，Wang X，Zhu T L. Maneuver detection methods for space objects based on dynamical model[J]. Advance in Space Research，2021，68(1)：71-84.

[10] Vaseghi S V. Advanced digital signal processing and noise reduction[M]. 4th Edition. New Jersey：John Wiley & Sons，Ltd，2008.

[11] 管洪杰，姚志成，刘岩，等. 基于容积卡尔曼滤波(CKF) 算法的卫星轨道实时预报方法[J]. 科学技术与工程，2013，13(32)：9622-9625.
Guan H J，Yao Z C，Liu Y，et al. Satellite orbit real-time prediction method based on Cubature Kalman Filter(CKF)algorithm[J]. Science Technology and Engineering，2013，13(32)：9622-9625.

[12] 唐宇，张伟，李星星，等. 基于机器学习的低轨卫星高精度轨道预报[J]. 武汉大学学报(信息科学版)，2025，50(11)：2373-2380.
Tang Y，Zhang W，Li X X，et al. High-precision orbit prediction for low Earth orbit satellites based on machine learning[J]. Geomatics and Information Science of Wuhan University，2025，50(11)：2373-2380.

[13] Bizalion H，Guillot A，Petit a，et al. Systematic TLE data improvement by neural network for most cataloged resident space objects[J]. Advances in Space Research，2023，72(7)：2649-2659.

[14] 崔文，郭超，张炜. 基于 TLE 双行根数的航天器机动检测方法[J]. 空间科学学报，2024，44(1)：151-158.
Cui W，Guo C，Zhang W. Spacecraft maneuver detection method based on TLE two-line element set[J]. Chinese of Journal of Space Science，2024，44(1)：151-158.

[15] 张涛涛，白显宗，郝嘉，等. 基于预报偏差的 LEO 航天器轨道异常检测[J]. 中国空间科学技术，2012，32(5)：40-46，68.
Zhang T T，Bai X Z，Hao J，et al. LEO spacecraft orbit anomaly detection based on prediction deviation[J]. Chinese Space Science and Technology，2012，32(5)：40-46，68.

[16] 彭利军. 航天器轨道监视与异常检测技术研究[D]. 长沙：国防科学技术大学，2012.
Peng L J. Research on spacecraft orbit surveillance and anomaly detection technology[D]. Changsha：National University of Defense Technology，Changsha，2012.

[17] 刘雁雨，常志巧，何峰，等. 基于小波分析的GEO卫星异常识别研究[J]. 大地测量与地球动力学，2010，30(5)：73-76.
Liu Y Y，Chang Z Q，He F，et al. GEO satellite anomaly identification based on wavelet analysis[J]. Journal of Geodesy and Geodynamics，2010，30(5)：73-76.

[18] 龙洗，杨乐平，乔琛远，等. GEO 卫星机动识别的自适应深度学习方法[J]. 宇航学报，2025，46(1)：51-67.
Long X，Yang L P，Qiao C Y，et al. Adaptive deep learning method for GEO satellite maneuver recognition[J]. Journal of Astronaut. 2025，46(1)：51-67.

[19] Xue B，Chuan L，Xiao P，et al. Mining two-line element data to detect

orbital maneuver for satellite[J]. IEEE Access，2019，7：129537-129550.

[20] Anantha N，Battu B T，Pruthwik M. TLE-based A2C agent for terrestrial coverage orbital path planning[EB/OL]. (2025-8-14)[2026-6-20]. https://arxiv.org/pdf/2508.10872.

[21] Cipollone R，LeonzioT，Calabrò G，et al. An LSTM-based maneuver detection algorithm from satellites pattern of life[C]//Proceedings of 2023 IEEE 10th International Workshop on Metrology for AeroSpace (MetroAeroSpace). Milan，Italy：IEEE，2023.

[22] 李炳臻，刘克，顾佼佼，等. 卷积神经网络研究综述[J]. 计算机时代，2021(4)：8-12，17.
Li B Z，Liu K，Gu J J，et al. Review of convolutional neural network research[J]. Computer Era，2021(4)：8-12，17.

[23] Yu F，Koltun V. Multi-scale context aggregation by dilated convolutions[C]//Proceedings of International Conference on Learning Representations (ICLR). [S. l]：ICLR 2016.

[24] Kipf T N，Welling M. Semi-supervised classification with graph convolutional networks[C]//Proceedings of International Conference on Learning Representations(ICLR). [S. l]：ICLR 2017.

[25] Prechelt L. Early stopping—but when?[M]. Heidelberg：Springer，2012.

作者简介：

王俊杰(2002－)，男，硕士研究生，主要研究方向：卫星轨道机动识别。

地址：湖南省长沙市开福区德雅路109号(410000)

手机：15022990916

E-mail：2282573372@qq.com

潘晓刚(1979－)，男，教授，博士，主要研究方向：智能任务规划技术、智能优化算法技术。**本文通信作者**。

地址：湖南省长沙市开福区德雅路109号(410000)

手机：13787115398

E-mail：panxiaogang_nudt@163.com

Non-Cooperative Micro-Thrust Detection Method for Satellites Based on Spatiotemporal Graph Convolution

Wang Junjie[1]，Sun Bowen[2]，Pan Xiaogang[2]，Peng Yong[1]，Wang Jiongqi[1]

（1. School of Science，National University of Defense Technology，Changsha 410000，China；
2. School of Systems engineering，National University of Defense Technology，Changsha 410000，China）

Abstract：The problem that the weak thrust of non-cooperative satellites, owing to its small magnitude and long duration, is highly susceptible to being submerged in observation noise，making detection extremely difficult，was prevalent in key stages such as precise orbit determination of low-Earth-orbit satellites and orbit control of deep-space probes，including midcourse corrections and near-Moon braking. To address this problem，a spatio-temporal graph convolutional network（ST-GCN） model based on temporal learning and spatial learning was proposed in this paper. By constructing the ST-GCN model，convolutional neural network（CNN），graph neural network（GCN），and feature fusion are introduced，and network parameter strategies such as adaptive learning rate scheduling and dynamic graph adjacency weights were adopted to improve detection accuracy and effectively solve the difficulty in detection caused by the extremely small magnitude of satellite micro-thrust. Experiments on three groups of low Earth orbit（LEO） satellite micro-thrust simulation datasets show that compared with current popular temporal models，the ST-GCN model proposed in this paper demonstrates the best overall performance in terms of both accuracy and efficiency. This method has the potential to be transferred to deep-space exploration scenarios and can provide a technical reference for the situational awareness of non-cooperative space targets.

Keywords：micro-thrust detection；spatio-temporal graph convolutional network（ST-GCN）；Non-cooperative satellite；spatio-temporal modeling；feature fusion

Highlights：

- Apply ST-GCN transfer learning to the detection of low-Earth orbit satellite micro-thrusts，while modeling both the temporal and spatial dependencies of orbital data.
- Construct a chain-like graph structure with time steps as nodes and adjacent points as edges，aligning with the Markovian characteristics of satellite orbits to achieve model lightweighting while ensuring feature expressiveness.
- Design a serial framework combining a dilated one-dimensional CNN and a residual GCN to effectively capture weak long-term temporal features.
- Employ adaptive learning rate scheduling and dynamic graph adjacency weighting strategies to improve model convergence speed and detection stability while reducing the risk of overfitting.

[责任编辑：杨晓燕，英文审校：宋利辉]