

基于深度不确定性估计网络的低轨卫星互联网故障预测方法

孙文宇, 张伟嘉, 王立民

(中国电科网络通信研究院, 河北 石家庄 050081)

摘要: 低轨卫星互联网技术具有创新性强、已有案例少且卫星数量众多、拓扑变化频繁、载荷类型繁杂等特点, 其故障类别难以靠专家系统或者工程经验等方法遍历和训练。同时传统的基于单颗卫星、单一功能的故障检测和诊断方法, 对于复杂多变条件下的不确定性故障难以进行预测。针对以上问题, 提出一种基于深度不确定性估计网络的低轨卫星互联网星座故障预测算法模型, 并基于地面星地联试获得的数据对所提方法进行验证。实验结果表明, 提出的方法可以提高已知故障的检测准确率, 尤其是对未知故障具有更准确的预测能力。

关键词: 低轨卫星互联网; 低轨卫星星座; 故障预测; 深度神经网络; 不确定性估计

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.2096-8930.2022025

Fault Detection Method of Low-Orbit Satellite Internet Based on Deep Uncertainty Estimation Network

SUN Wenyu, ZHANG Weijia, WANG Limin

Network and Communication Research Institute of CETC, Shijiazhuang 050081, China

Abstract: Low-orbit satellite internet technology is a highly innovative direction of development with few existing cases. Low-orbit satellite internet systems are characterised by a large number of constituent satellites, frequent topology changes, and complex payload types. The fault categories of such complex systems are difficult to be completely mastered with methods such as expert systems or practical engineering experience. Moreover, traditional fault detection and diagnosis methods tailored for a single satellite and a single functionality are difficult to predict uncertain faults under sophisticated and transient environment conditions. For the above problems, a fault prediction model based on the deep uncertainty estimation network for low-orbit satellite internet was proposed, and was verified used self-collected data from simulated joint satellite-ground tests. Experimental results showed that the proposed method could improved the detection accuracy of known types of faults and demonstrated particular effectiveness in predicting faults of unknown types.

Keywords: low-orbit satellite internet, low-orbit satellite constellation, fault detection, deep neural network, uncertainty estimation

0 引言

低轨卫星互联网是当前全球发展的重点, 与以往的低轨星座相比, 低轨卫星互联网具有跨代差别: 一是由窄带移动功能向频段更高、速率更高、容量更大的宽带通信功能发展; 二是卫星规模呈数十倍、百倍增长, 典型的窄带移动低轨星座仅有几十颗卫星, 而典型卫星互联网星座则有几百至数万颗卫星; 三是可能发生故障的层级从物理层、链路层向网络层、传输层乃至应用层延展; 四是由单一通信网络功能向兼容导航增强、遥感增

强、ADS-B、AIS等多功能拓展; 五是由支持单一空中接口体制向适应更多垂直行业, 支持多样化空中接口体制方向发展^[1-4]。总体上看, 系统复杂度大幅增加, 很多场景以及系统之间的关联性没有先例, 这必然导致系统故障类型预测以及故障本身预测的难度增加。当前传统的故障诊断方法有: 基于专家系统^[5]和基于测试性模型^[6-7]的故障诊断技术。基于专家系统的故障诊断技术需要依赖大量的专家经验, 且知识库覆盖的故障模式有限, 不能对未出现的故障进行诊断。基于测试性模型的故障诊断方法,

需要对系统（或设备）组成、故障机理、因果逻辑等关系有深入了解，并建立精确的物理或者数学模型。这些缺点都阻碍了相关故障诊断技术的进一步应用。在此背景下，数据驱动的故障诊断技术表现出优异的性能，其以机器学习算法为基本理论，利用数据挖掘技术进行故障特征学习和提取，建立故障特征和故障模式之间的联系，从而达到检测与诊断的目的，具有良好的可扩展性、可迁移性且无人工先验依赖性，因此其成为该领域研究的热点。

1 系统故障不可预测性分析及数据驱动故障诊断方法

1.1 卫星互联网星座故障不可预测性分析

传统的高轨卫星和低轨移动星座系统，主要有单设备故障，比如卫星平台故障、卫星天线故障、卫星射频故障等。对于低轨卫星互联网而言，故障涉及面大大增加，不仅扩展至链路层、网络层，具有很多相互关联的系统级不可预测故障以及系统动态性导致的多场景的不可预测故障。如图1所示，单设备故障相对增加星上处理设备（比如卫星路由设备、卫星网控设备等）的故障，链路层故障涉及同轨星间、异轨星间、星地馈电、星地不同频段的用户链路等多种类型，特别是，由于用户站或终端对卫星不断切换、关口站对卫星和卫星链路不断切换，因此这一类链路有很多不同的场景，导致链路故障的不可预测性。

系统级故障更是复杂，且具有更大概率率的不可预测性，比如波束控制策略与波位预期的符合性出现问题，最

佳路由关联、多星跨南北极切换，多空中接口并存干扰、非通信载荷控制等多个方面均存在故障的不确定性。

1.2 数据驱动的故障诊断方法

基于数据驱动的故障诊断方法结合统计分析法、信号处理方法及人工智能方法等领域的理论和方法。统计分析法主要有主成分分析法、核主成分分析法，隐马尔可夫模型。信号处理方法主要有小波变换、希尔伯特变换。人工智能方法主要有支持向量机和神经网络法，在国内外已取得了一定的成果。例如，参考文献[8]利用支持向量机的方法实现卫星姿态控制系统执行器故障的检测与隔离，同时利用小波变换提升了支持向量机的泛化能力以及寻优效率。参考文献[9]提出一种基于神经网络的鲁棒故障检测与隔离方法，构造两个递归神经网络分别用于识别敏感器和执行器故障。参考文献[10]针对传统方法难以实现对卫星姿态控制系统进行故障分离的问题，提出一种使用随机森林算法的卫星姿态控制系统的故障诊断方法。参考文献[11-12]结合长短期记忆（Long Short Term Memory, LSTM）网络和时序特征处理卫星时序故障预测问题。参考文献[13]提出一种残差收缩网络，将软阈值函数引入深度残差网络的内部，提高从强噪声信号学习特征的能力，并得到更高的故障检测准确率。

上述基于数据驱动的卫星故障检测算法大多是针对系统的单点故障预测，其分类结果（模型输出）都是确定性的。如图2所示，如果模型输入的样本是未知故障数据时，期望模型可以给出不确定的预测而非某具体的分类结果。在卫星故障预测的高敏感性领域，模型算法

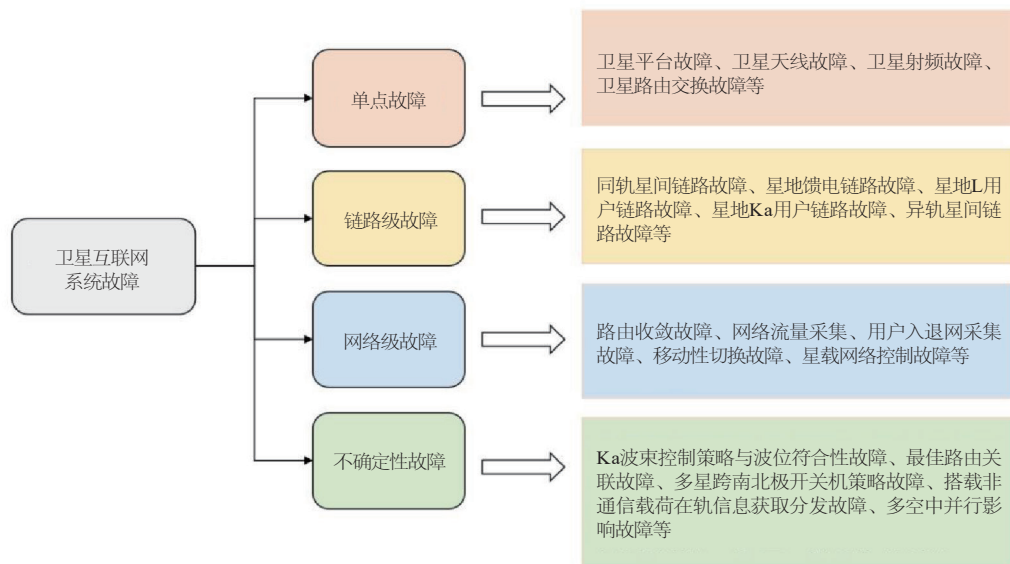


图1 低轨卫星互联网星座主要故障类型示意

结果的可靠性和可解释性尤为重要，而不确定性估计方法恰好能够提高模型的可靠性和鲁棒性，并具有对模型“未知”的数据进行判断的能力。不确定性估计最早应用于机器学习中，主要通过给出模型输出的概率分布或置信度值的方式表示模型对输入样本类别或预测回归值的确信程度。例如，参考文献[14]利用不确定性对输入模型的异常点数据进行检测。参考文献[15]尝试利用不确定性的性质和估计方法处理基于风险的决策问题以及可靠性估计。

基于上述背景，本文提出将不确定性估计方法与深度神经网络模型相结合的方法，使模型能够对于自身能力之外的输入样本，也就是未知故障状态给予较低的置信度，即赋予模型“知道自己不知道”的能力。本文的主要工作与贡献如下。

(1) 提出了一种基于深度不确定性估计网络的故障检测方法，相比支持向量机（Support Vector Machine, SVM）^[16]、随机森林^[10]、基准深度神经网络^[17]等算法进一步提高了故障分类的准确性。

(2) 首次在卫星故障检测中引入不确定性估计。通过在特征嵌入中引入不确定性分量，并利用模型对引入不确定性分量的特征（均值）和不确定性（方差）进行学习，赋予了模型对“未知故障”检测的能力。

(3) 优化了面向故障检测的深度模型损失函数。该损失函数可同时对不确定性和特征进行约束，既解决了不确定性分支训练难以收敛的问题，也避免了训练中的偏向性和平凡解问题。

2 深度不确定性估计网络架构与理论

本节对基于深度不确定性估计网络的卫星互联网故障检测模型及相关理论进行分析。首先定义卫星故障检

测数据中的不确定性，对这种不确定性进行特征分布建模，并描述如何将不确定性引入特征向量。随后对该网络的主干网络、特征提取网络 and 不确定性分支进行详细分析，并进一步讨论不确定性分支的约束和损失函数。深度不确定性估计网络整体的模型结构示意如图3所示。

2.1 不确定性估计

首先假设输入样本的特征嵌入为服从高斯分布的隐空间采样得到的随机性嵌入。故障类别相同的样本都是围绕类别特征中心分布上的一个点，定义不确定性为每条数据到特征中心的距离，远离特征中心的数据定义为不确定性数据，靠近特征中心的数据定义为确定性数据。这样得到的特征嵌入形式如下

$$P(F_i|x_i) = \mathcal{N}(F_i | \text{GAP}(\tilde{f}), \sigma_i^2) \quad (1)$$

其中， F 代表输入样本 x 的特征向量， GAP 代表全局平均池化， $\tilde{f}$ 是主干网络最后一层输出的特征。分布 $\mathcal{N}$ 是对角多元高斯分布，并且高斯分布的参数均值与方差均是通过网络前向传播的输出值。模型输出的原始特征 $f = \text{GAP}(\tilde{f})$ 作为均值，以参数为 Ω 的全连接层的预测值 $\sigma_i = \text{FC}_\Omega(f)$ 的平方作为方差。通过对原始特征 f 引入不确定性分量，得到不确定性的特征 f' ，具体形式如下

$$f'_i = f_i + \mu_i \quad (2)$$

其中， f_i 代表的是卫星采集数据的第 i 维特征嵌入，引入的不确定性分量是 μ_i 。在训练阶段输入数据 x_i 得到的特征嵌入 F 是引入不确定性分量后的 f' ，而在测试阶段输入数据 x_i 得到的特征嵌入 F 为原始特征 f 。

此外，在进行不确定性损失函数训练时，也会对不确定性分量进行额外的约束，一般的约束手段是计算某种分布的距离度量，并消减源分布与目标分布间的度量大小。本方法中使用KL散度量两个概率分布的距离，

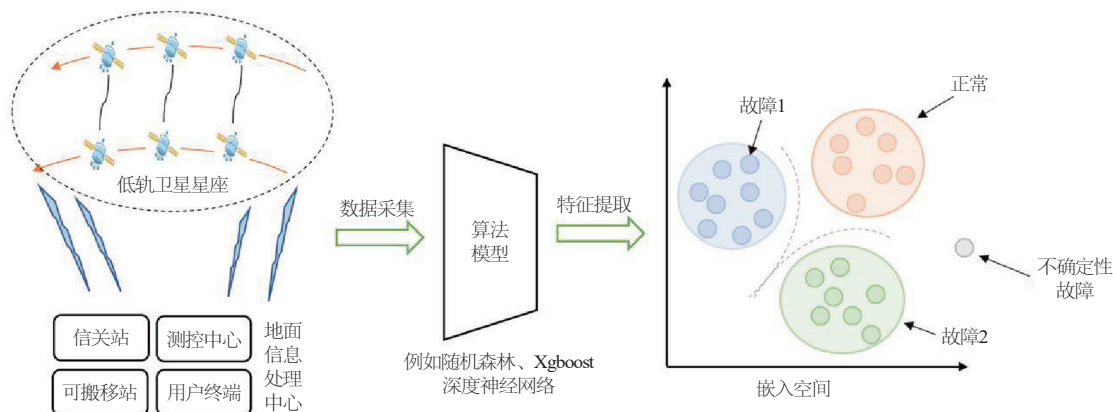


图2 低轨卫星互联网星座中的未知故障问题

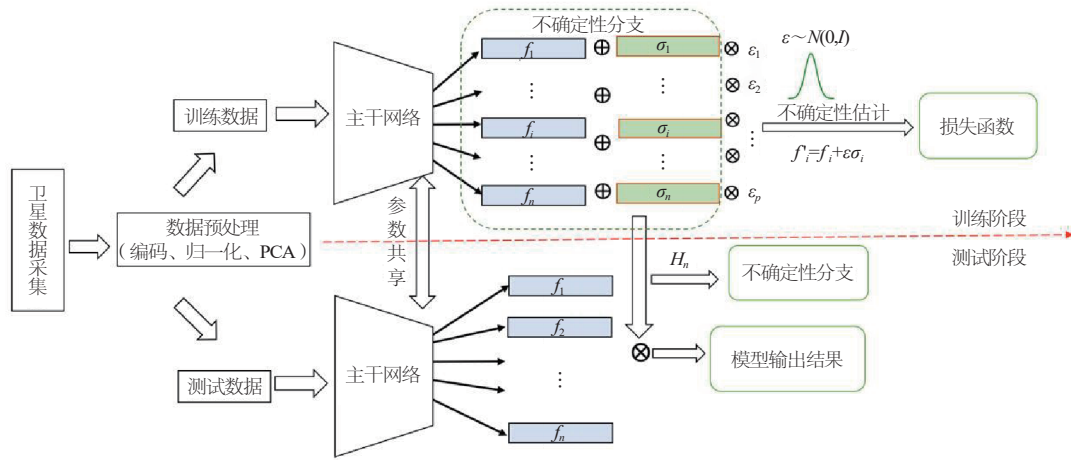


图3 深度不确定性估计网络模型结构示意图

其本质上是计算使用分布 P_2 来编码分布 P_1 所需的额外平均位数，具有非对称性，其定义如下

$$\text{KL}(P_1\|P_2) = \sum_{x \in X} P_1(x) \log \frac{P_1(x)}{P_2(x)} \quad (3)$$

其中， P_1 代表实际卫星采集数据特征分布， P_2 代表引入不确定性的数据特征分布。除KL散度外，也可以使用JS散度或Wasserstein距离等距离度量作为故障特征的不确定性约束项。以单机的板卡温度、电压为例，采集值超出阈值范围或骤变为确定性故障，测量值出现阈值范围内的非正常规律性为不确定性故障。

2.2 主干网络与不确定性分支结构

本小节主要分析深度不确定性估计网络模型的实现细节，包括主干网络和不确定性分支结构。进行特征提取的主干网络使用了典型的深度神经网络（DNN）架构。为了更多关注和比较不确定性估计模块带来的提升，忽略基础主干网络带来的性能影响，本文应用主流的残差网络ResNet18^[13]作为深度不确定估计网络的主干特征提取网络。残差模块的输出为 $H(x)=F(x)+x$ ，即垂直分支和跳跃连接的加和。其中 $F(x)$ 就是中间卷积层部分的残差。由于近路连接对 x 的求导结果等于1，因此可以反传本层的梯度结果至上一层，这样能够缓解链式梯度法则引起的梯度消失现象。通过残差模块可以使网络的层数叠加到比一般深度神经网络更深，从而具备更强的特征提取能力。在本文中由于输入的卫星采集数据是一维的，因此特征图和卷积核的高度 H 均为1，卷积核大小 K 等于通道数 C 。主干网络中的残差模块如图4所示。

需特别注意的是，本文应用的主干网络可被替换为其他网络，例如VGG^[18]、Inception^[19]和ShuffleNet^[20]等卷

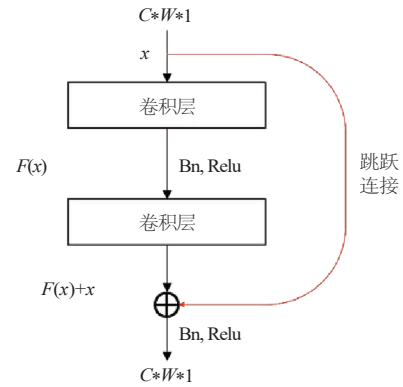


图4 主干网络中的残差模块

积神经网络。

深度不确定性估计网络模型中不确定性分支的网络结构示意图如图5所示。将经过预处理后的卫星采集的数据送入模型主干网络提取特征后，原始特征嵌入 f 对特征图进行全局平均池化操作。不同于正常的网络结构（一般由几个全连接层构成），在训练阶段将原始特征嵌入直接通过颈部网络计算分类损失函数并进行分类训练。该网络并不直接使用原始特征 f ，而是通过添加不确定性分量的特征嵌入 f' 进行训练。此外，本文还在模型的颈部网络结构中添加了BNNeck结构，纠正了不同样本特征之间的偏离程度且不影响特征嵌入自身的判别性与可分性，最大程度同时利用分类损失和不确定性损失各自的贡献。

通过对原始的数据特征 $f=\text{GAP}(\tilde{f})$ 添加不确定性分支，得到不确定性的特征向量 f' 。如果将不确定性分支的输出向量 σ 作为不确定性估计值，并使用重参数化技巧（Reparameterisation）对 σ 的每一维进行采样，这样得到 f' 的

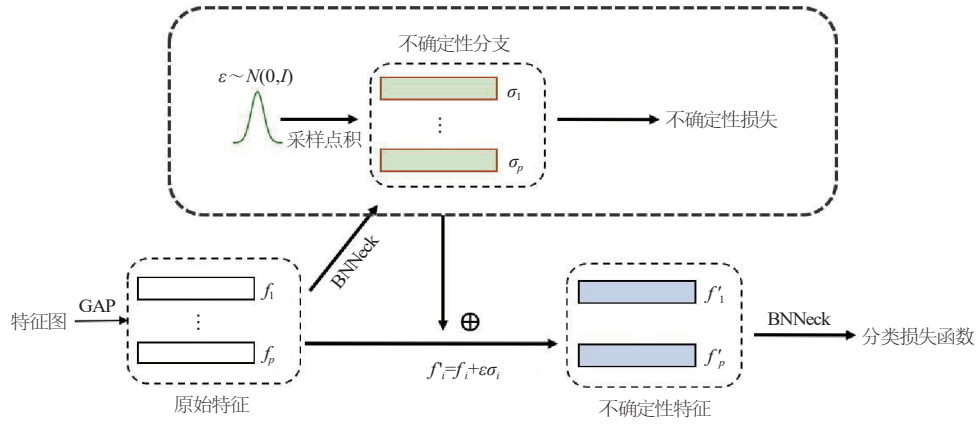


图5 深度不确定性估计网络模型中不确定性分支的网络结构示意

表达式为

$$f'_i = f_i + \epsilon \sigma_i \quad (4)$$

其中，引入的不确定性分支是 $\epsilon * \sigma$ ， ϵ 是通过标准正态分布 $\epsilon \sim N(0, I)$ 采样得到的， σ_i 是模型的输出。这里采样得到 ϵ 的方式，借鉴了变分自编码器VAE^[21]中编码后计算特征向量的方法。采样操作具有不可导性，导致反向传播无法计算梯度，本文通过对标准正态分布的采样值直接乘以 σ ，在不确定性分支中使用重参数化技巧，来避免直接采样所带来的不可导性。原始特征对应的不确定性估计 $\sigma \in R^{C^0 \times 1}$ 由不确定性分支的输出得到，计算方式如下

$$\sigma^2 = \text{FC}(\text{GAP}(\tilde{f}); \Omega_\sigma) \quad (5)$$

其中， $\text{FC}(*; \Omega_\sigma)$ 代表不确定性分支的全连接层，采用参数为 $\Omega_\sigma \in R^{C^0 \times C^0}$ 的全连接层结构。将不确定性分支的全连接层输出的每一维 σ_i^2 求调和平均数，就能够计算出样本整体的不确定性估计值 $\hat{\sigma}^2 \in R$ 。事实上，这里的不确定性估计值其实是确信度，表示网络对其自身预测的样本特征的置信程度，也就是说预测值越大，代表不确定程度越小。之所以这样表示是为了便于计算并且与其他不确定性估计方法一致。

$$\frac{1}{\hat{\sigma}^2} = \sum_{i=1}^{C^0} \frac{1}{\sigma_i^2} \quad (6)$$

2.3 损失函数优化

卫星故障检测任务可以视为一个分类问题，其基本思想是，最大化类间方差和最小化类内方差，同时还要对不确定性进行约束。基于此，本文整体的损失函数 L_{SU} 由两部分组成。 L_{SU} 的具体计算方式如下

$$L_{\text{SU}} = \frac{1}{N} \sum_{i=1}^N \left[(1 + \sigma_i^2) (\alpha_1 L_{\text{Softmax}}) + \alpha_2 U_{\text{KL}} \right] \quad (7)$$

其中，第一部分 L_{Softmax} 代表使用交叉熵损失函数作为故障类别信息损失函数。 L_{Softmax} 也可以被替换为其他分类损失，因为分类损失函数不是本文研究重点，所以这里仅使用SoftmaxLoss作为损失函数， L_{Softmax} 损失函数的具体表达式如下

$$L_{\text{Softmax}} = \frac{1}{N} \sum_i \left(-\log \left(\frac{e^{w_{y_i} f'_i}}{\sum_c e^{w_c f'_i}} \right) \right) \quad (8)$$

第二部分是不确定性约束的正则项 $U(\sigma)$ ，用来避免训练阶段出现平凡解或训练坍塌的现象。如果模型对预测的卫星采集样本很确信，则网络输出的 σ^2 趋于1，此时不确定性损失对整体损失没有贡献，仅存在分类损失。模型对其预测的卫星采集样本最不确定时，网络输出的 σ^2 趋于0，这种情况下分类损失函数的大小基本不对参数的更新产生影响。如果没有不确定性的正则项约束，会导致不论输入样本如何，不确定性分支都输出 $\sigma^2=0$ ，即不确定性损失函数值一直为0，不确定性分支无法得到训练。为了防止训练中出现平凡解，损失函数中需加入 $U(\sigma)$ ，如果使用KL散度对不确定性项进行约束，得到 $U_{\text{KL}}(\sigma)$ 的具体表达式为

$$U_{\text{KL}}(\sigma) = \text{KL} \left[N(f'_i | f_i, \sigma_i^2) \| N(\epsilon | 0, I) \right] = -\frac{1}{2} \sum_i \left(\log \sigma^2 + 1 - |f|^2 - \sigma^2 \right) \quad (9)$$

其中， $U_{\text{KL}}(\sigma)$ 代表通过KL散度对不确定性项进行约束的损失函数。通过Sigmoid层将不确定性分量 σ 的值域映射到[0,1]区间，因此 $U(\sigma)$ 在值域范围单调递减，正则项约束损失函数起到避免过拟合的效果。如果 σ 偏小，则分类损失函数将减小而不确定性损失将增大；如果 σ 偏大，则分类损失函数将增大而不确定性损失会减小，因此避免了训练阶段出现的偏向性问题。除此之外，还能够进一步

将不确定性项的约束替换为JS散度。使用JS散度的正则项 $U_{js}(\sigma)$ 的具体形式如下

$$U_{js}(\sigma) = JS\left[N\left(f_i|f_i, \sigma_i^2\right) \parallel N(\varepsilon|0, I)\right] = -\frac{1}{4} \sum_i \left(\log \frac{4\sigma^2}{(\sigma^2+1)^2} + \frac{|f|^2}{(\sigma^2+1)} \right) \tag{10}$$

使用JS散度能够解决分布间非对称的问题，并且由于其值域相对可控，在训练阶段更有利于搜索不确定性项约束权重的相对大小，并达到更好的约束结果。同时，与KL散度进行约束相比较而言，过早地引入不确定性项约束会导致训练坍塌现象；JS散度可以解决不确定性分支训练无法收敛的问题，因此无须对分类损失进行预训练。

3 仿真实验

3.1 数据与实验细节

本文以自研低轨卫星模拟中心机接收载荷单电路板卡的遥测数据为例，通过人为设置正常设备状态和控制板卡、调制解调板卡、星间交换、星地交换等4种载荷设备的故障状态，每种状态各接收90 min遥测数据，建立实验数据集。选择其中的正常状态和3种故障状态数据对深度不确定性估计模型进行训练，将没有参与训练的1种设备故障状态（故障状态4）作为未知故障（模型未进行学习）进行测试。数据集状态统计信息见表1。

表1 卫星遥测数据集状态统计信息

遥测数据状态	数据量/条	是否参与训练
正常状态	12 124	是
故障状态1	5 312	是
故障状态2	5 769	是
故障状态3	4 793	是
故障状态4	4 902	否

本文实验代码全部基于Python v3.6.7和PyTorch1.4.0完成，对所有训练和测试的原始数据进行了one-hot编码和归一化以及PCA降维操作，同时对训练数据的每一维引入5%大小的高斯噪声进行数据增强。为了消除在训练模型中各种随机因素的影响，保证对比实验的有效性与合理性，本文使用Kaiming初始化^[22]进行参数初始化并在训练模型时手动设置了随机数种子，确保在所有实验中模型的初始化参数保持一致。

本文的方法在实验中设置batch大小为32，特征向量的维度被固定为128维，选择ResNet18作为模型主干网

络。式(7)中的参数 α_1 、 α_2 通过实验对比分别被设置为1、0.01。训练阶段设置优化函数为Adam，同时加入 1×10^{-5} 的权重约束，并使用大小为0.9的Nesterov动量。迭代次数（Epoch）被设置为60，通过 T 等于20的周期余弦退火算法对学习率进行变化。训练共分为3个阶段，首先只进行Backbone和分类模块的训练；之后冻结Backbone和分类模块参数，只进行不确定性模块训练；最后解冻所有参数进行微调，并选取第三阶段所有Epoch中最优的模型。Backbone的初始学习率设为0.001，不确定性估计分支与分类分支的初始学习率设为0.01。

3.2 实验结果与分析

为了对比提出的深度不确定性估计网络卫星故障预测算法，本文通过系统采集的遥测数据与常见的故障分类算法做了对比验证实验。不同算法模型对已知故障的预测结果见表2，其中对比方法SVM类型被设置为C_SVC、Sigmoid核，随机森林的参数被设置为 $k=200$ 、 $m=5$ ，基准CNN模型使用ResNet18作为主干网络、Softmax作为分类损失函数，训练过程与参数细节均与本文算法保持一致。

表2 不同算法模型对已知故障的预测结果

算法模型	正常状态	故障状态1	故障状态2	故障状态3
SVM	99.13%	88.64%	90.25%	96.42%
随机森林	99.17%	96.78%	98.36%	98.29%
ResNet18+Softmax	99.47%	98.89%	99.04%	99.15%
ResNet18+ U_{kl} (本文方法)	99.63%	99.26%	99.40%	99.51%
ResNet18+ U_{js} (本文方法)	99.70%	99.58%	99.67%	99.27%

从表2的实验结果中能够观察到，本文提出的不确定性估计模型在已知的故障检测数据集中，与SVM、随机森林、基准CNN模型（ResNet18）相比，在准确率指标上均有明显提升。其中，使用深度学习网络的方法相比传统机器学习方法提升较多，且使用本文方法的效果更为显著，性能提高了近3%（99.58% vs 96.78%，故障状态1）。 U_{kl} 和 U_{js} 分别是在KL散度和JS散度约束下的不确定性约束项，从表2中也可以看出，采用JS散度要明显优于采用KL散度对不确定性约束项的效果。但是以上对比方法无法对未知故障进行预测，原因是它们确定了分类的类别数量。

为了能够对比其他方法对未知故障的检测能力，本文在训练时将任务作为二分类任务进行训练，即只分为正常和故障两类。在实际应用中，SVM、随机森林、基准CNN模型等分类器一般是将任务作为多分类任务进行

训练，并不具有对故障状态分类的能力，只能对训练时已有的具体故障状态进行预测。这也是以上对比方法在实际卫星故障检测应用的缺点。本文方法可以通过判断不确定性分支的输出对未知故障状态进行判断，通过设定阈值判断不确定估计值的大小，来判断输入样本的特征分布是否产生偏移，以此识别未知故障。

表3 不同算法模型对未知故障的预测结果

算法模型	正常状态	故障状态1~3	故障状态4
SVM	99.16%	91.76%	56.57%
随机森林	99.21%	97.85%	60.72%
ResNet18+Softmax	99.32%	99.26%	59.22%
ResNet18+U _{kl} （本文方法）	99.28%	99.04%	98.14%
ResNet18+U _{js} （本文方法）	99.10%	99.21%	98.76%

不同算法模型对未知故障的预测结果见表3，对于模型已知状态的数据结果与表2中的实验结果类似。同样使用ResNet18作为主干网络提取特征，深度不确定性估计网络的故障检测效果更显著，也验证了本文方法的有效性。同时，可以明显看出在未知故障状态（故障状态4）下，本文方法可以通过不确定性估计对未知故障给出准确判断，原因在于网络的不确定性分支能够对样本中的异常进行精准的学习和判断。对比方法SVM、随机森林和基准CNN模型等对故障状态4预测准确率仅略高于50%，说明传统的点估计的预测模型无法对未知故障给出准确预测。

卫星通信系统与环境的复杂性导致了卫星遥测数据可能带有噪声，而不确定性估计方法一般能够提升网络对噪声数据的可靠性与鲁棒性。为了验证本文提出的

深度不确定性估计网络模型是否提升了噪声数据的鲁棒性，本文在实验中添加了在高斯噪声影响下的噪声样本后对模型进行测试。不同噪声干扰对算法模型预测结果的影响如图6所示，可以看出数据中的噪声越大，本文方法的效果越明显，通过加强对噪声样本的约束，使模型能够对异常数据提取更具判别性的特征，因此提升了算法对噪声样本的抗噪能力。

表4 不确定性项约束权重对实验的影响

约束权重	F1-score
$\alpha_2/\alpha_1=0$	99.08%
$\alpha_2/\alpha_1=0.001$	99.19%
$\alpha_2/\alpha_1=0.01$	99.58%
$\alpha_2/\alpha_1=0.1$	97.34%
$\alpha_2/\alpha_1=1$	82.55%

通过设置大小不一样的不确定性项约束权重对比实验，验证了本文超参数和对模型的影响程度，具体见表4。从表4中能够看出， α_2/α_1 的大小直接影响算法对不确定性数据的敏感程度，为方便起见令 $\alpha_1=1$ 。当 α_2 被设置为0.01时算法的F1-score最高，所以本文算法中的不确定性项约束权重大小被设为0.01。对比表4中的实验结果能够发现，若 α_2/α_1 的值过小则约束损失函数不具有意义，不确定性分支估计分数都接近0，此时相当于直接使用深度神经网络模型进行训练。若 α_2/α_1 的值过大，则不确定性约束损失函数相比故障分类损失会提供过大的梯度，使得算法过分倾向对数据偏移的学习而减少对故障分类判别性特征的学习，导致算法带有严重偏向性且整体性能大幅降低。

4 结束语

本文针对卫星故障任务中存在未知故障的问题，提出了一种基于深度不确定估计网络的故障预测算法。通过对低轨卫星遥测数据特征不确定性建模分析和在遥测数据特征嵌入中引入不确定性分量，同时通过模型对不确定性分量的特征（均值）和不确定性（方差）进行学习，能够在提高已知故障分类准确性的同时具备对未知故障预测的能力。此外该方法也优化了模型损失函数，通过对不确定性和特征进行约束，解决了不确定性分支训练难以收敛的问题。

面对基于数据驱动的卫星故障任务，还有以下几个方向值得研究。

（1）随着卫星设备老化，卫星数据的特征可能发生概念漂移现象，如何通过无监督的学习方法预测和纠正

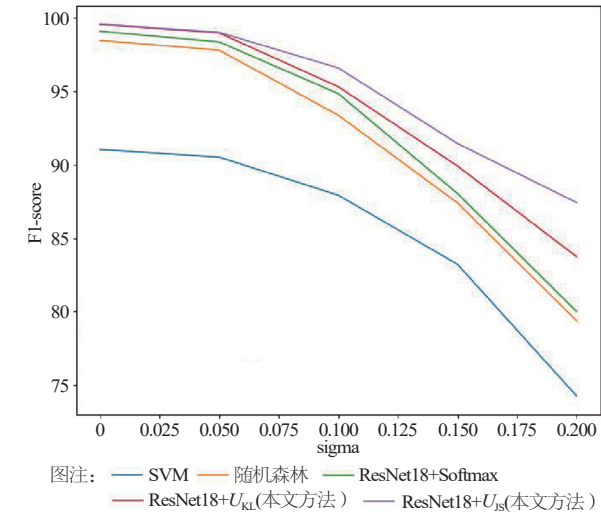


图6 不同噪声干扰对算法模型预测结果的影响（F1-score）

这种现象现在还未有深入研究。

(2) 可以尝试使用贝叶斯神经网络作为本文方法的主干网络, 模型中所有的参数都服从各自的分布, 这样网络能够直接输出概率分布的预测, 其关键在于高频次的采样操作带来更高的计算复杂度, 导致贝叶斯神经网络难以进行训练。

(3) 与基于专家知识的故障检测方法共同构建混合模型, 将模型输出的不确定性故障进行故障树分析, 提高整体系统的故障检测性能。

参考文献:

- [1] 吴巍. 天地一体化信息网络发展综述[J]. 天地一体化信息网络, 2020, 1(1): 1-16.
WU W. Survey on the development of space-integrated-ground information network[J]. Space-Integrated-Ground Information Networks, 2020, 1(1): 1-16.
- [2] 肖永伟, 孙晨华, 赵伟松. 低轨通信星座发展的思考[J]. 国际太空, 2018(11): 24-32.
XIAO Y W, SUN C H, ZHAO W S. Discussion on the problem of LEO communication constellation system design[J]. Space International, 2018(11): 24-32.
- [3] 何辞, 张亚生, 孙晨华, 等. 低轨星座组网及地面IP路由技术适应性分析[J]. 天地一体化信息网络, 2020, 1(1): 36-41.
HE C, ZHANG Y S, SUN C H, et al. Analysis on low-earth-orbit constellation networking and adaptability of ground IP routing technology[J]. Space-Integrated-Ground Information Networks, 2020, 1(1): 36-41.
- [4] 孙晨华, 章劲松, 赵伟松, 等. 高低轨宽带卫星通信系统特点对比分析[J]. 无线电通信技术, 2020, 46(5): 505-510.
SUN C H, ZHANG J S, ZHAO W S, et al. Comparative analysis of GEO and LEO broadband satellite communication system[J]. Radio Communications Technology, 2020, 46(5): 505-510.
- [5] TAKAKI R, HASHIMOTO M, HONDA H, et al. ISACS-DOC: automatic monitoring and diagnostic system for scientific satellite(space development and AI)[J]. Journal of Japanese Society for Artificial Intelligence, 2006, 21.
- [6] ZHANG K, JIANG B, SHI P. Adaptive observer-based fault diagnosis with application to satellite attitude control systems[J]. Second International Conference on Innovative Computing, Information and Control (ICICIC 2007), 2008, 4(8): 1921-1929.
- [7] BALDI P, BLANKE M, CASTALDI P, et al. Combined geometric and neural network approach to generic fault diagnosis in satellite actuators and sensors[J]. IFAC-PapersOnLine, 2016, 49(17): 432-437.
- [8] 陈辛, 魏炳翌, 闻新. 基于支持向量机的卫星执行机构故障诊断研究[J]. 中国空间科学技术, 2018, 38(2): 47-55.
CHEN X, WEI B Y, WEN X. Study of support vector machine based faults diagnosis for satellite's actuators[J]. Chinese Space Science and Technology, 2018, 38(2): 47-55.
- [9] TALEBI H A, KHORASANI K, TAFAZOLI S. A recurrent neural-network-based sensor and actuator fault detection and isolation for nonlinear systems with application to the satellite's attitude control subsystem[J]. IEEE Transactions on Neural Networks, 2009, 20(1): 45-60.
- [10] 黄瑾, 刘洋, 钟麦英, 等. 利用随机森林算法的卫星控制系统故障诊断[J]. 宇航学报, 2021, 42(4): 513-521.
HUANG J, LIU Y, ZHONG M Y, et al. Fault diagnosis of satellite attitude control systems using random forest algorithm[J]. Journal of Astronautics, 2021, 42(4): 513-521.
- [11] 钱昭勇, 曹裕华, 张雷. 基于CNN-LSTM的在轨卫星故障预测分析[J]. 军事运筹与系统工程, 2021, 35(1): 57-63, 72.
QIAN Z Y, CAO Y H, ZHANG L. Fault prediction and analysis of the on-orbit satellite based on CNN-LSTM[J]. Military Operations Research and Systems Engineering, 2021, 35(1): 57-63, 72.
- [12] 董静怡, 庞景月, 彭宇, 等. 集成LSTM的航天器遥测数据异常检测方法[J]. 仪器仪表学报, 2019, 40(7): 22-29.
DONG J Y, PANG J Y, PENG Y, et al. Spacecraft telemetry data anomaly detection method based on ensemble LSTM[J]. Chinese Journal of Scientific Instrument, 2019, 40(7): 22-29.
- [13] ZHAO M H, ZHONG S S, FU X Y, et al. Deep residual shrinkage networks for fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2020, 16(7): 4681-4690.
- [14] DEVRIES T, TAYLOR G W. Learning confidence for out-of-distribution detection in neural networks[J]. arXiv preprint arXiv:1802.04865, 2018.
- [15] KIUREGHIAN A D, DITLEVSEN O. Aleatory or epistemic? Does it matter? [J]. Structural Safety, 2009, 31(2): 105-112.
- [16] CHEN K Y, CHEN L S, CHEN M C, et al. Using SVM based method for equipment fault detection in a thermal power plant[J]. Computers in Industry, 2011, 62(1): 42-50.
- [17] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [19] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4,

inception-resnet and the impact of residual connections on learning[C]//Thirty-first AAAI conference on artificial intelligence. [S.l.:s.n.], 2017.

- [20] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6848-6856.
- [21] KINGMA D P, WELING M. Auto-encoding variational bayes[J]. arXiv preprint arXiv:1312.6114, 2013.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1026-1034.

[作者简介]



孙文宇（1996—），男，硕士，中国电科网络通信研究院工程师，主要研究方向为低轨卫星互联网、卫星通信运行控制、深度学习。



张伟嘉（1994—），男，硕士，中国电科网络通信研究院工程师，主要研究方向为深度学习与机器学习、卫星通信系统。



王立民（1976—），男，中国电科网络通信研究院高级工程师，主要研究方向为卫星有效载荷、低轨卫星互联网。

2022 11.23—25

北京·国家会议中心



关注官方公众号

中国国际信息通信展览会（PT EXPO）是经国务院批准，由工业和信息化部主办的ICT行业盛会。自1990年创展以来，已经成功举办三十届。展会集“产学研用”四位一体，融展示、交流、合作、共享于一炉，通过展示最新信息通信行业发展成果，向世界输出“中国好方案”，是我国乃至全世界优秀的信息通信成果展示平台。

展示面积



40000+

展商



400+

政企团组



450+

演讲嘉宾



800+

论坛



50+

媒体



3000+

专业观众



180000人次

成果发布



20+



扫码沟通

- **权威：**工业和信息化部唯一主办的信息通信行业展会
- **高端：**跨界融合，与各行业领军人面对面交流，为行业发声
- **高效：**紧扣行业需求，同期举办多场商洽会、产品发布会、产融对接会等，搭建直面交流平台
- **高曝光：**数千家国内外媒体人齐聚现场，对展会进行持续性曝光
- **多成果：**各行联手，行业白皮书、最新科研成果、重大合作等争相亮相展会现场

收录声明

本刊对发表的文章，拥有出版电子版、网络版版权，并拥有和其他网站交换信息的权利。本刊支付的稿酬中已经包含上述费用。

Space-Integrated-Gound Information Networks has the copyright to publish electronic edition, online edition of the published articles, and has the right to exchange information with other sites. The expenses have been included in the fee paid by editorial department.

道德声明

本刊发表的论文是作者独立取得的原创性研究成果，无一稿多投；论文内容不涉及国家机密；未曾以任何形式用任何文种在国内外公开发表过；论文内容不侵犯他人著作权和其他权利。若发生一稿多投、侵权、泄密等问题，论文作者将承担全部责任。

The authors of *Space-Integrated-Gound Information Networks* guarantee that their submitted articles are original and contain nothing confidential. The said article is only submitted to *Space-Integrated-Gound Information Networks*. The said article has not been published before and has not been submitted elsewhere for print or electronic publication consideration. The said article is no way whatever a violation or an infringement of any existing copyright or license from the third party. Otherwise, the authors of the said article shall take the blame for the violation or infringement of the related copyright and the leakage of secrets.