

2022-02-03T21:00

《大气科学》杂志编辑部：

你们好！

我是国防科技大学气象海洋学院大气科学专业的硕士研究生，罗文杰，研究方向是人工智能方法在卫星遥感中的应用。这篇《利用改进的 GoogLeNet 深度学习模型识别 COSMIC-2 掩星信号中的反射信号》的论文是由我和我导师项杰教授以及杜华栋副教授共同完成的，我们希望能在此发表。

论文的主要内容：介绍了掩星反射信号的研究现状和无线电全息技术；提出一种基于改进的 GoogLeNet 深度学习模型，并应用于 COSMIC-2 掩星探测数据来识别反射信号；分析了反射信号与掩星数据质量的关系。

论文选择《利用改进的 GoogLeNet 深度学习模型识别 COSMIC-2 掩星信号中的反射信号》这个题目进行研究，有以下几个重要的意义：

(1) GPS 掩星探测技术作为一种新型的大气遥感探测手段，已成为大气科学领域研究与应用热点，我们选择了最先进的 COSMIC-2 掩星资料进行研究，相较于 COSMIC、CHAMP 和 FY-3C GNOS 等掩星资料将具有更好的效果，同时也是对 COSMIC-2 掩星数据质量的检验；

(2) 关于掩星反射信号识别、掩星反射信号所包含的大气信息等问题的研究是掩星资料处理与应用的重要研究内容，国外许多学者对此已开展了持续而深入的研究，而国内很少有人进行掩星反射信号相关的研究，而我们的研究弥补了国内关于掩星反射信号研究的空白；

(3) 提出了改进的 GoogLeNet 深度学习模型用来识别掩星中的反射信号，相较于无线电掩星气象卫星应用中心 (ROM SAF) 所使用的支持向量机 (SVM) 方法效果显著提高，推进了深度学习在掩星资料处理上的应用。

(4) 准确识别掩星反射信号，研究反射信号中包含的大气信息对于掩星资料同化，提高数值天气预报精度具有重要的意义。

论文的创新点在于：

(1) 采用了最新的 COSMIC-2 掩星数据提取掩星中的反射信号，检验了 COSMIC-2 掩星数据的良好质量；

(2) 提出了改进的 GoogLeNet 深度学习模型用来识别掩星中的反射信号；

(3) 初步分析了反射信号与掩星数据质量的关系，检验了前人的结论。

据了解，贵刊主要报道大气科学领域的创新性研究成果，是国内大气科学领域的顶级刊物，我们一直重点关注，贵刊所载的卫星气象学、大气遥感等方向的诸篇大作，我都仔细拜读过，受益颇多。我们的论文将卫星气象学、掩星资料应用和深度学习等多个领域融合在一起，主题和内容符合贵刊的要求。我们通过检索了解到，贵刊刊登 GPS 掩星相关的论文相对较少，近期刊发了两篇关于云海-2 掩星资料同化的文章，而与本文主题类似的文章最近在贵刊还没有涉及，如果我们的论文有幸被录用，希望能够补充和丰富贵刊在掩星资料相关领域的研究。

非常感谢您审阅本论文，期待早日收到您的审稿意见。若对本论文有任何疑问，请及时与我联系。

顺致

编安！

投稿人：罗文杰

2021 年 6 月 6 日

利用改进的 GoogLeNet 深度学习模型识别 COSMIC-2 掩星信号中的反射信号

罗文杰 项杰 杜华栋

国防科技大学气象海洋学院, 长沙 410008

摘要 GPS 掩星探测技术作为一种先进的大气探测手段,已广泛用于数值天气预报,气候和空间天气研究。掩星探测存在的问题之一是容易受到地球表面反射信号的干扰,识别和分离掩星探测信号中的反射信号有助于将掩星数据同化到数值天气预报系统中去,具有重要意义。本文提出一种基于改进的 GoogLeNet (Im-GNet)深度学习模型,并应用于 COSMIC-2 掩星探测数据来识别反射信号。本文选择了 2020 年 1 月 1 日至 9 日的 COSMIC-2 掩星数据(conPhs 文件),进行质量控制后,利用无线电全息方法得到掩星信号的无线电全息功率谱密度图像,然后训练得到 Im-GNet 深度学习模型,Im-GNet 模型测试的准确率达到 96.4%,显著高于支持向量机(SVM)方法的结果。本文还分析了反射信号对掩星数据的影响,掩星事件的地理分布以及掩星反演数据(atmPrf 文件)与 NCEP 12 小时预报值(avnPrf 文件)的折射率比较表明:有反射信号的掩星事件数据质量更好,所包含的大气信息更丰富。

关键词 COSMIC-2 掩星 深度学习 无线电全息技术 GoogLeNet

Identify the reflected signal in the COSMIC-2 occultation signal using the improved GoogLeNet deep learning model

LUO Wenjie, XIANG Jie, Du Huadong

College of Meteorology and oceanography, National University of Defense Technology,
ChangSha 410008

Abstract GPS occultation detection technology, as an advanced atmospheric detection method, has been widely used in numerical weather forecasting, climate and space weather research. One of the problems in occultation detection is that it is easily interfered by the reflected signals on the earth's surface. Identifying and separating the reflected signals in the occultation detection signal helps to assimilate the occultation data into the numerical weather prediction system, which has important significance. This paper proposes a deep learning model based on improved GoogLeNet (Im-GNet) and applies it to COSMIC-2 occultation detection data to identify reflected signals. This article selects the COSMIC-2 occultation data (conPhs file) from January 1 to 9, 2020. After quality control, the radio holography method is used to obtain the spatial spectrum image of the occultation signal, and the Im-GNet deep learning model is trained, The accuracy rate of Im-GNet model test reached 96.4%, which is significantly higher than the result of support vector machine (SVM) method. This paper also analyzes the impact of reflected signals on occultation data. The geographic distribution of occultation events, and the refractivity comparison between the occultation inversion data (atmPrf file) and the NCEP (National Centers for Environmental Prediction) 12 hour forecast files (avnPrf file) shows that the quality of the occultation event data with reflection signals is better and the atmospheric information contained is richer.

Keyword COSMIC-2 occultation, Deep Learning, Radio Holographic technology, GoogLeNet

作者简介 罗文杰,男,1997 年出生,硕士研究生,研究方向是人工智能方法在卫星遥感中的应用。E-mail: 835321654@qq.com

通讯作者 项杰,教授,主要从事反问题及资料同化等方面的研究。E-mail: xjieah@aliyun.com

资助项目 国家自然科学基金资助项目 41475021

1 引言

GPS 掩星探测技术作为一种新型的大气遥感探测手段，具有高精度、高垂直分辨率、全球准均匀覆盖、全天候观测、长期稳定等特点（Kursinski et al.,1996; Kuo et al.,2004），已广泛用于数值天气预报，气候和空间天气研究。如，我国台湾与美国联合实施的掩星探测项目“气象、电离层和气候星座观测系统（COSMIC，2006-2011）”成为最成功的掩星探测业务应用系统。在 COSMIC 项目之后，台湾和美国又推出了名为 FORMOSAT-7/COSMIC-2 的后续掩星探测项目。COSMIC-2 的六颗卫星于 2019 年 6 月 25 日成功发射到赤道附近的低倾角（ 24° ）轨道；7 月 16 日卫星上的 3 个仪器（主载荷 3G 掩星系统（Tri-GNSS Radio-occultation System, TGRS）、副载荷离子速度计（Ion Velocity Meter）和射频信标机（Radio Frequency Beacon））开始激活。COSMIC-2 团队在经过仪器性能评估和算法优化后，于 2019 年 12 月 10 日开始发布掩星探测数据，数据的起始日期为 2019 年 10 月 1 日。由于应用了先进的 TGRS 接收机和高增益天线，结合算法的改进，COSMIC-2 掩星探测信号具有前所未有的高信噪比，这使得 COSMIC-2 项目每天能获得 4000 多条高质量的中低纬度（ $45^\circ\text{S}\sim 45^\circ\text{N}$ ）中性大气廓线（弯角、折射率、温度、压力和湿度），这些数据具有高准确度、高精度和高垂直分辨率等特点（Schreiner et al.,2020）。

人们在分析 GPS 无线电掩星数据时发现掩星数据中存在 GPS 信号与地球表面相互作用的信息（Beyerle and Hocke,2001; Beyerle et al.,2002）。图 1 所示为掩星探测示意图；除了常见的主体折射信号（通常称为“直接信号”）外，还经常可以检测到另一个信号：该信号具有较小的频率偏移（由于传播介质的变化和多普勒频移效应）和较低的仰角，通常表现为多路径传播、超折射（异常大的折射率梯度）和强衰减等畸变现象之一，而且有时这个信号并不太容易确定。研究表明，这个信号是来自地球表面的反射信号，常被跟踪直接信号的接收机无意捕获。为了表述的方便，此信号被称为“反射信号”（Aparicio et al.,2018）。

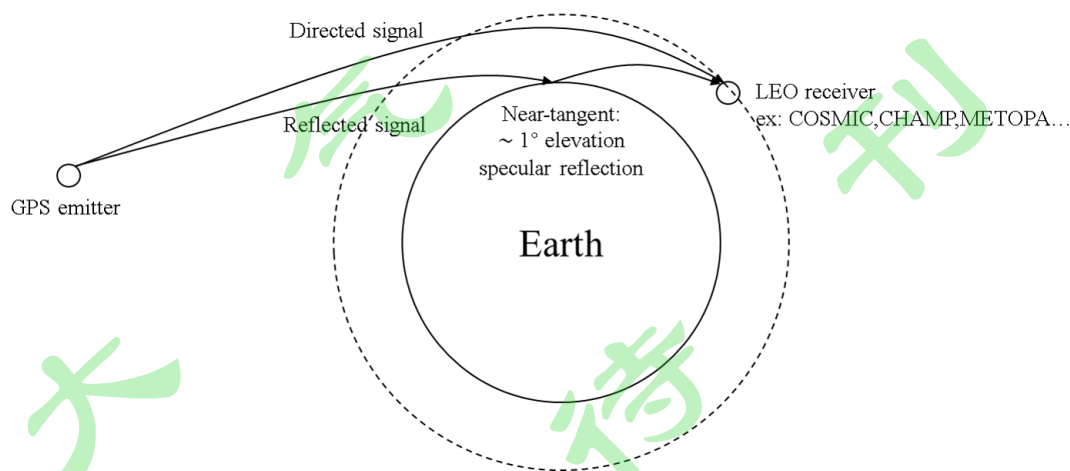


图 1 GNSS-RO 的直接信号与反射信号

Fig. 1 Direct signal and reflected signal of GNSS-RO (Global Navigation Satellite System - Radio occultation)

反射信号与低层大气和地表存在相互作用, 因此, 除了由直接信号提供大气信息之外, 反射信号也成为确定低层大气和地表性质的潜在的信息源 (Aparicio et al., 2018)。如, Beyerle et al. (2002) 发现反射分量中信噪比 (SNR) 值可能与海面特征 (如有效波高或波浪坡度统计) 存在联系; Cardellach et al. (2004) 利用来自反射信号的干涉相位数据对极区冰川高度进行了尝试性测量; Boniface et al. (2011) 表明, 无线电掩星事件中的反射信号可能包含有关对流层最低层 (最低几公里) 的地球物理信息, 并提供低对流层状态的信息, 因此是气象信息的潜在来源, 同时还提出: 对直接和反射无线电信号的频移进行分析可以帮助确定大气或表面特性。

在数值天气预报 (NWP) 应用上, 有研究工作表明 (Cardellach et al., 2008), GPS 掩星信号中存在反射信号是掩星数据质量良好的指标, 这样的掩星探测廓线与 ECMWF 分析场更一致。Cardellach et al. (2008) 注意到, 具有反射信号的掩星事件从直接信号中获得的大气折射率廓线要比没有清晰反射信号的掩星事件更接近于 NWP 模型的估计值, 同时使用 ECMWF 的数据进行了测试。Aparicio et al. (2018) 表明具有强反射信号的掩星事件的数据与 NWP 结果有良好的先验一致性, 并尝试提取可用于 NWP 背景下的大气信息, 特别是提供最低 5 公里内温度和湿度廓线的附加信息 (仅限于海洋反射事件)。

据统计, 大约 40~60% 的掩星信号包含来自大气和地球表面的反射信号。然而, 在反演掩星弯角的过程中, 这些受到“污染”的掩星观测值在反演过程中通常被视为异常值而被抛弃 (尤其是在波动光学方法中) (Hu et al., 2018)。

为了处理无线电掩星 (RO) 信号的多路径传播、衍射等情形而发展的基于傅立叶积分算子的技术也同样能够反演反射信号的弯角廓线 (Gorbunov et al., 2018)。这些技术包括: 基于傅立叶积分算子 (FIO) 的正则变换 (CT) (Gorbunov, 2001), 全谱反演 (FSI) (Jensen, 2003) 和相位匹配 (PM) (Jensen et al., 2004), 以及第二类正则变换 (CT2) (Gorbunov and Lauritsen, 2004)。其中 FSI 和 CT2 方法在数值上最有效 (Gorbunov et al., 2006)。

目前, 识别反射信号的手段主要是滑动窗口的无线电全息技术 (又称为滑动频谱方法) (Hocke et al., 1999; Beyerle and Hocke, 2001; Cardellach et al., 2008; Boniface et al., 2011), 该技术已用于提取反射场 (Aparicio et al., 2018)。此外也可以使用基于 Wigner 分布函数 (WDF) 的技术来识别反射信号 (Gorbunov et al., 2016)。Beyerle et al. (2002) 利用几何射线追踪和多相位屏模拟方法定量分析了无线电全息图中的特征频移变化, 这些特征频率变化主要与信号在地表反射点的海拔高度和地面的折射率有关; 他们用于检测反射信号的主要方法是通过无线电全息图的目视检查。Cardellach et al. (2008) 提出了基于支持向量机 (SVM) 的无线电全息方法, 用于快速自动检测 RO 事件中的反射信号, 然而, 该方法在识别反射信号的特定范围方面仍然有局限性, 并且需要大量迭代才能收敛。为了缩小错误识别的范围并提高计算效率, Hu et al. (2018) 提出了基于梯度下降人工神经网络的无线电全息方法 (GDANN-RH), 以提高检测成功率并减少计算时间。

除了以上的识别反射信号的工作以外，还有一些人开展了反射信号的反演工作。如 Boniface et al. (2011) 首先利用无线电全息滤波方法从 GPS 掩星信号中分离出反射信号，然后利用一种扰动方法进行反演，确定了最低层大气对反射信号特性的相关贡献，并提取了对直接传播信号反演解的补充信息。Gorbunov et al. (2018) 提出了基于修正的 CT 方法的反演算法，并进行了反演试验。

鉴于 Cardellach et al. (2008) 和 Hu et al. (2018) 使用的都是 COSMIC 数据，而最新的 COSMIC-2 中性大气掩星探测数据已经对外公开发布，可以下载使用，因此，本文旨在把无线电全息技术应用于 COSMIC-2 掩星探测数据以识别出其中的反射信号；通过建立基于改进的 GoogLeNet 反射信号分类模型，获得更好的反射信号自动检测效果。此外，本文还将分析反射信号与掩星数据质量的关系。本文的结构如下：第二节介绍空间谱工具的算法原理，描述反射信号的局地空间谱特征；第三节通过改进 GoogLeNet 网络结构建立反射信号分类模型，并与 SVM 方法进行对比分析；第四节讨论反射信号与掩星资料的偏差；第五节为总结讨论。

2 掩星信号局地空间谱的提取

2.1 无线电全息技术

本文中掩星信号局地空间谱的提取采用无线电全息方法。设低轨 (LEO) 卫星测得的电磁波场为：

$$u(t) = A(t) \exp[ik\Psi(t)] \quad (1)$$

其中， $A(t)$ 是振幅， k 是波数， $\Psi(t)$ 是相位， t 是时间。选定参考相位：

$$u_m(t) = \exp[ik\Psi_m(t)] \quad (2)$$

其中， $\Psi_m(t)$ 可以通过对观测波场 $u(t)$ 的相位 $\Psi(t)$ 进行平滑得到。那么，

$$\Delta u(t) \equiv \frac{u(t)}{u_m(t)} = A(t) \exp[ik(\Psi(t) - \Psi_m(t))] \quad (3)$$

称为无线电全息图 (radiohologram) (Hocke et al., 1999)。

对无线电全息图(3)进行滑动窗的傅里叶变换 (Sokolovskiy, 2001) 得到：

$$c(t, \omega) = \frac{1}{2\pi} \int_{t-T/2}^{t+T/2} \Delta u(t') \exp(-i\omega t') dt' \quad (4)$$

其中， t 是时间窗口的中心， T 是窗口的范围大小， ω 是局地的频率。谱 $|c(t, \omega)|^2$ 取最大值时的频率 ω_m 代表 t 时刻的射线。谱 $|c(t, \omega)|^2$ 的变化情况可以在 $t-\omega$ 图 (时-频图) 上表示出来。同时，谱 $|c(t, \omega)|^2$ 的变化情况也可以在影响参数-弯角图上表示出来 (Gorbunov et al., 2000; Gorbunov, 2002)。

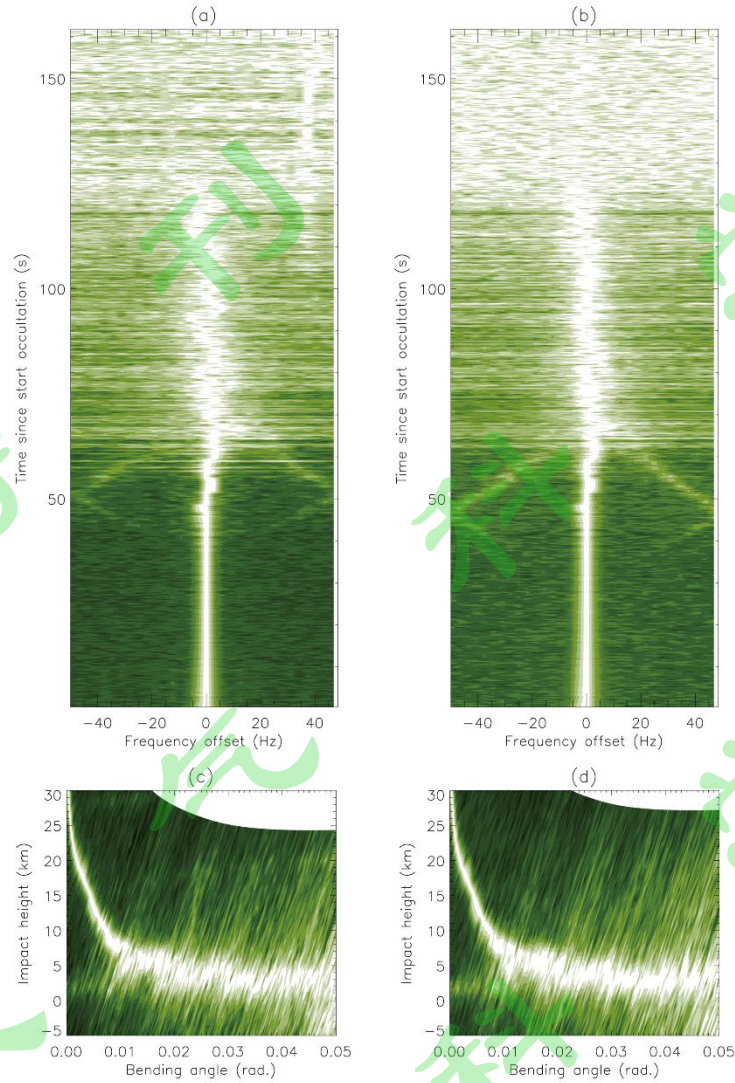


图 2 一次 COSMIC-2 掩星事件（掩星编号为 C2E1.2020.001.03.50.G19，纬度 1.75°N，经度 54.02°E）的空间谱图像。（a）和（b）分别为 L1 和 L2 信号的时间-频率域空间谱；（c）和（d）分别为 L1 和 L2 信号的影响高度-弯角域空间谱

Fig. 2 A spatial spectrum image of a COSMIC-2 occultation event (occultation ID is C2E1.2020.001.03.50.G19, latitude 1.75°N, longitude 54.02°E). (a) and (b) are the time-frequency domain spatial spectrum of the L1 and L2 signals, respectively; (c) and (d) are the impact height and bending angle domain spatial spectrum of the L1 and L2 signals, respectively

图 2 给出了采用无线电全息方法计算 COSMIC-2 掩星事件的空间谱图像（时间-频率域与影响参数-弯角域）的一次示例，本文的研究仅使用到 L1 信号的时间-频率域空间谱，影响高度-弯角域空间谱则在后续的掩星资料质量控制与误差分析的研究中起到重要作用。

2.2 反射信号的视觉识别

Beyerle and Hocke（2001）提出了第一个证据，即 GPS/MET 无线电掩星实验中观察到的 GPS 信号包含反射信号分量，他们采用无线电全息方法来分离直接和反射信号成分。

GNSS 掩星数据中存在反射信号的特征体现在其局地空间谱图像中, 反射信号的频谱不同于主要的直接 (折射) 信号的频谱; 该频谱呈牛角形, 随着时间的流逝, 它逐渐趋向直接信号的中心频率 (出现在下行掩星事件中, 对于上行掩星事件则是偏离) (Cardellach et al., 2008)。这种结构通常很清晰, 而且易于通过人工视觉识别。

本文使用 COSMIC-2 掩星数据进行研究, 借助无线电全息技术, 通过确定局地空间谱图像中的特征频移来检测反射信号。本文处理的数据仅限于下行掩星 L1 波段 (1.57542 GHz) 的 C/A 信号, 原因是上行掩星的反射信号容易出现偏离, 且 L2 波段 (1.22760 GHz) 的发射功率较低, 信号容易受到噪声的影响。

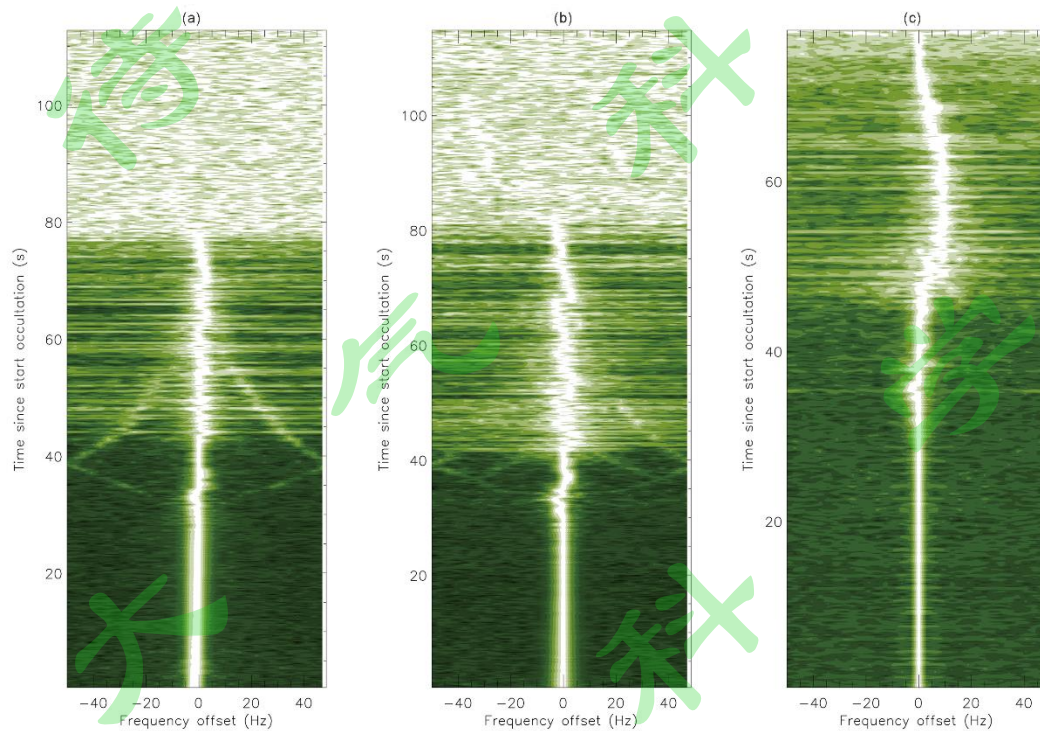


图 3 COSMIC-2 掩星 L1 信号局地空间谱图像 (a) 清晰的反射信号 (b) 模糊的反射信号 (c) 无反射信号

Fig. 3 Local spatial spectrum image of COSMIC-2 occultation L1 signal (a) clear reflection signal (b) unclear reflection signal (c) no-reflection signal

图 3 给出了三种类型的掩星信号的局地空间谱图像。可以看到, 在图 3a 中, 由于直接信号的功率占主体, 因此, 直接信号以接近零的频率出现在局地空间谱图像中, 并受大气和电离层中的小尺度结构的影响而加宽。38~65s 之间出现在直接信号两端的较低强度 (振幅) 分量则是反射信号, 该信号的频率随时间而变化, 并趋向于直接信号的中心频率 (零频率附近, 在 60s 之后与直接信号逐渐会聚。图中较亮的水平横线是信号混叠造成的 (由于 COSMIC-2 信号的采样率为 100Hz, 频率在 50 Hz 以上的信号会发生混叠)。在大约 80s 之后, 掩星事件结束, 直接信号消失。而由于电离层 (总电子含量 TEC 起主要作用)、中性大气层 (水汽的影响最大) 和地球表面 (复杂的粗糙度产生散射) 对反射信号传播的影响, 同时卫星仪器性能或影响到反射信号的接收和处理, 图 3b 则在 35~50s 之间出现了模糊的反射

信号特征；图 3c 是典型的无反射信号示例：50s 之前直接信号清晰而集中，50~80s 由于信号干扰致使直接信号发生小幅度的频移和散焦。

3 改进的 GoogLeNet 反射信号分类模型

GoogLeNet 是 Christian Szegedy 等人在 2014 年提出的一种全新的深度学习结构，它和 LeNet-5、AlexNet、VGG Net、ResNet 等其它经典 CNN 结构一样通常应用于图像处理领域。GoogLeNet 卷积神经网络在 2014 年的 ImageNet 分类任务中击败 VGG (Visual Geometry Group) 夺得冠军，以往 AlexNet、VGG 等结构都是通过增大网络的深度（层数）来获得更好的训练效果，但增加网络层数的会带来 overfit、梯度消失、梯度爆炸等负作用。GoogLeNet 提出了一个全新的深度 CNN 架构 Inception 代替了单纯的“卷积+激活”的传统操作，达到提高性能而又不大量增加计算量的目的，把对卷积神经网络的研究推上新的高度。本文选择 GoogLeNet 的 Inception (3a、3b、4a) 模块作为神经网络主体，以改进 GoogLeNet 模型用于反射信号的识别。

3.1 数据及预处理

我们从 [https://data.cosmic.ucar.edu/gnss-ro/cosmic2/nrt/\[2020-04-06\]](https://data.cosmic.ucar.edu/gnss-ro/cosmic2/nrt/[2020-04-06]) 下载 2020 年 1 月 1 日至 9 日的 COSMIC-2 掩星近实时 conPhs 数据 (Level-1b)，进行质量控制后，利用无线电全息方法得到 17826 张下行掩星 L1 信号的空间谱图像。通过专家标记得到有清晰反射信号的图像 5903 张，反射信号模糊的图像 4519 张和无反射信号的图像 7404 张。在实验中，删除了反射信号模糊的图像，仅对有清晰反射信号和无反射的图像进行分类识别。

将上述 13307 张实验图像合理截取成 224*224 像素的灰度图像并进行归一化。随机选择总数据 70% 的数据作为训练数据集，15% 的数据作为验证数据集，剩余 15% 的数据作为测试数据集，并保存为 pickle 型数据，将大幅减少模型训练前的数据集导入时间。

为了评估所提出方法的性能，将其与 SVM 方法进行比较，设置以下三种方案进行测试：

- (1) 改进的 GoogLeNet 模型 (Im-GNet)；
- (2) 线性核函数的 SVM (Li-SVM)；
- (3) 高斯核函数的 SVM (Ga-SVM)。

为了评估每种方案的运算时间，表 1 给出了实验的详细配置。其中，使用 Tensorflow 框架来建立和训练 Im-GNet 模型，SVM 不是研究的重点，则直接调用 Scikit-learn 库来实现。

表 1 实验计算机的软硬件配置

Table 1 The configuration for hardware and software

项目	配置
硬件	内存 (RAM): 16G 处理器: Inter(R)Core(TM)i7-9850H CPU
操作系统	Windows 10 专业版
计算机语言	Python 3.7.0
机器学习	GoogLeNet: Tensorflow 2.1.0

3.2 改进的 GoogLeNet 网络结构

GoogLeNet 模型最大的特点在于引入了 Inception 模块。以图 4 中 Inception 3a 模块为例，该模块共有 4 个分支，左边第一个分支对输入进行 1×1 卷积，它可以跨通道组织信息，提高网络的表达能力；第二个分支先使用了 1×1 卷积，然后连接 3×3 卷积，相当于进行了两次特征变换；第三个分支类似，先是 1×1 的卷积，然后连接 5×5 卷积；最后一个分支则是 3×3 最大池化后直接使用 1×1 卷积。Inception 模块的引入会显著提高参数的利用效率，其原因在于：卷积层提升表达能力主要依靠的是增加输出通道数，但会产生计算量增大和过拟合的副作用（杨真真等,2018）。

经过多次实验，我们保留 GoogLeNet 卷积神经网络 Inception（3a、3b、4a）作为神经网络主体，并在前 2 个 Inception 结构后添加 BatchNorm Layer 保持训练稳定。在输出结果前增加 2 个 Dense Layer 与 1 个 Dropout Layer，提升输出空间的描述能力同时防止过拟合，具体网络结构见图 4。

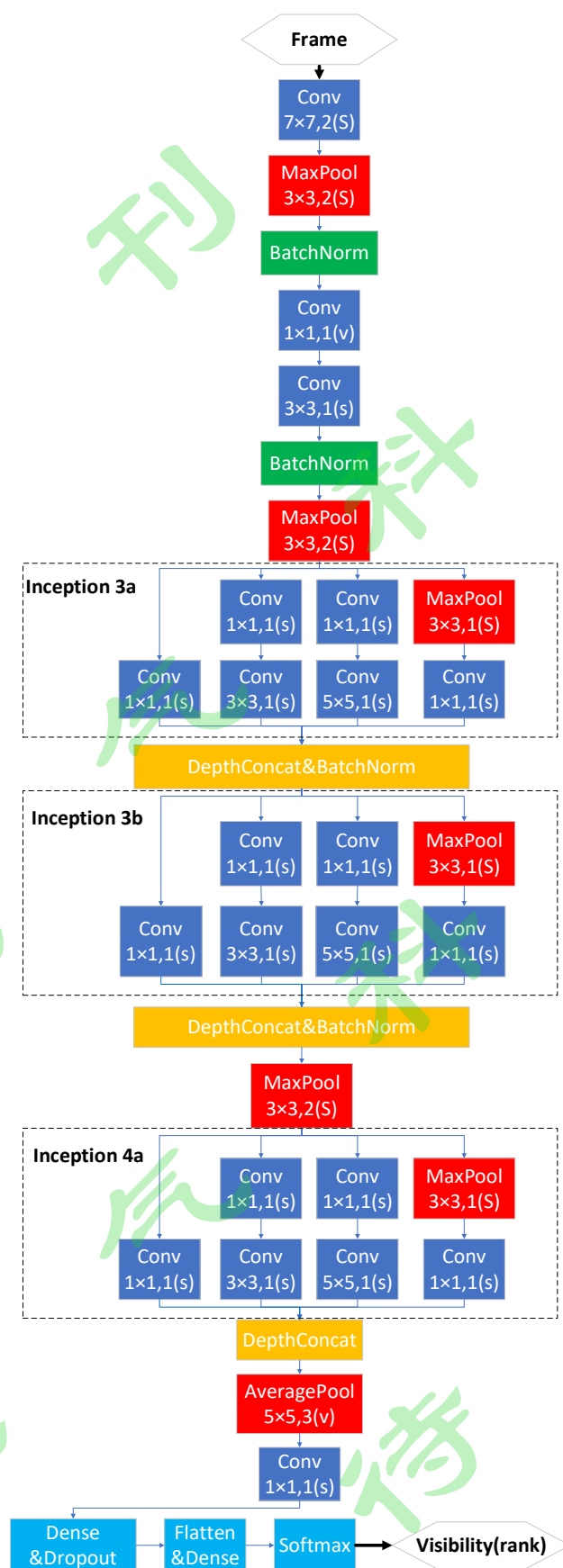


图 4 Im-GNet 网络结构

Fig. 4 Im-GNet network structure

改进的 GoogLeNet 基于 Tensorflow 2.1.0 架构搭建网络，采用了交叉熵损失函数（Cross-entropy Loss Function），主要参数设置如下：Shuffle（Buffer_size）=12000，Batch_size=50，Epochs=100，Learning_rate=0.0001，其中设置每个阶段节点学习率（Learning_rate）减半并延续至下一个阶段。通常来说，合理的参数设置，可以提高神经网络的泛化性能和模型的训练速度，使得梯度下降的方向更加准确，优化模型使之达到最佳的性能。

从理论上来说，GoogLeNet 比 SVM 更加健全和完善，更有利于解决非线性问题，在图像分类上具有更优的性能，它的 Inception 结构既能保持网络结构的稀疏性，又能利用密集矩阵的高计算性能。下面给出具体的处理结果。

3.3 模型结果及评估

图 5 给出了 Im-GNet 模型训练过程中准确率和损失率的变化曲线。在第 25 次迭代之前，验证集的准确率和损失率曲线均存在较大波动，训练集的准确率呈上升趋势，损失率呈下降趋势，模型欠拟合。第 25 次迭代往后，验证集损失率逐渐上升，其他曲线趋于平稳，模型过拟合。在第 25 次迭代时，验证集的准确率达到 0.96，损失率为次低值且无大的波动，模型达到最佳拟合效果，即模型训练完成。之后，将测试集数据输入模型得到其准确率为 96.4%。

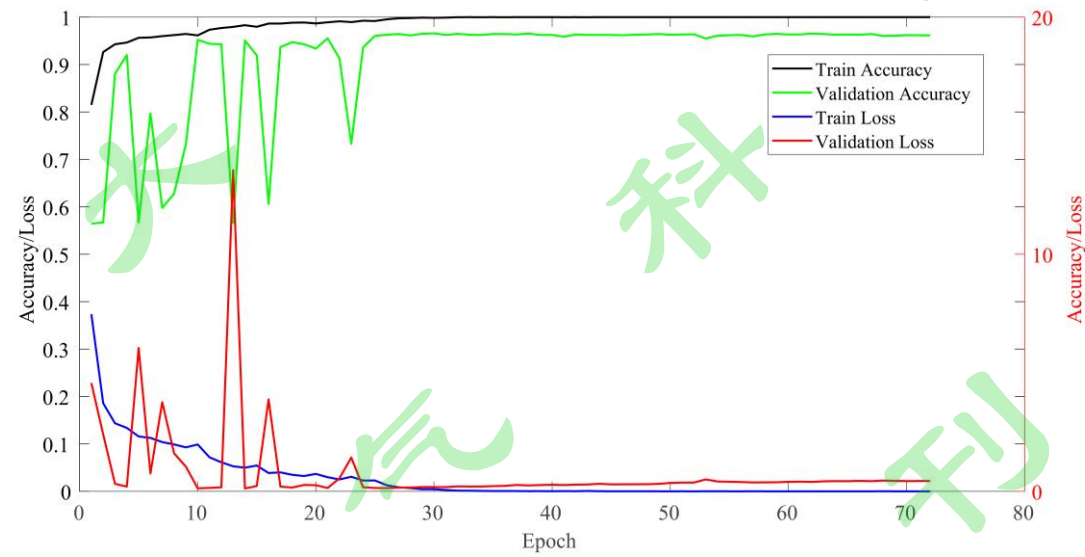


图 5 Im-GNet 模型的训练过程

Fig. 5 The training process of the Im-GNet model

表 2 给出了三种模型的测试结果。从表 2 可以看出，Im-GNet 的运算时间要远远高于另外两种模型，原因是 Inception 结构迭代的时间复杂度更高，但相对于 Li-SVM 和 Ga-SVM 来说反射信号的识别准确率分别提高了 15.5% 和 22.6%。这表明，Im-GNet 的掩星反射信号的识别效果要优于 SVM。

表 2 三种模型的测试结果

Table 2 Test results of the three models

方案	模型	模块/核函数	准确率 (%)	运算时间 (s)
1	Im-GNet	Inception (3a、3b、4a)	96.4	12423
2	Li-SVM	线性核函数	80.9	2053
3	Ga-SVM	高斯核函数	73.8	2804

4 反射信号对掩星数据的影响

关于反射信号对掩星数据的影响，一直以来有许多学者在进行研究。如 Beyerle et al. (2002) 分析了 2001 年 3 月 14 日至 6 月 10 日间的 3783 次 CHAMP 掩星事件的地理分布，观察到热带纬度地区出现掩星反射信号的数量显著低于其他纬度。Cardellach et al. (2008) 利用 COSMIC 掩星资料进行研究发现：GPS 掩星信号中存在反射信号是掩星数据质量良好的指标，包含反射信号的掩星数据廓线与同一时间、同一位置的 ECMWF 分析场更接近。此外，有研究表明 (Cardellach and Oliveras, 2016)：在海洋上，反射事件的比例随纬度增加；反射事件与粗糙度/风况不相关，因为在这些倾斜的几何形状中，有效粗糙度趋于零，而与海表温度有一定的相关性。

由于 COSMIC-2 掩星探测项目应用了先进的仪器并改进了算法，其数据具有更高的信噪比和精度，相较于使用 CHAMP、COSMIC 在中低纬度的掩星探测数据来说，本文使用 COSMIC-2 数据研究反射信号对掩星数据的影响应该具有更优的效果。

4.1 反射信号与掩星事件的地理分布

通过计算发现，本文所研究的数据集中约有 44.4% 的掩星事件呈现出反射信号特征；在呈现出反射信号的掩星事件中，有 95.6% 发生在海洋上，仅有 4.4% 发生在陆地上，且主要分布在湖泊或沙漠 (见图 6)。这表明：掩星事件的地理分布表明接收器捕获的反射信号的清晰度 (相干性) 与地表性质存在相关性，在相对光滑的表面 (海洋) 中尤其常见，这是准镜面反射的预期行为，可保持信号相干性。海洋和湖泊等均匀表面是掩星信号的良好反射器，陆地多为不规则的表面，通常会产生漫散射使信号失去相干性，故陆地掩星事件极少接收到反射信号。

然而，海洋上接近一半的掩星事件没有反射信号，这些掩星事件主要出现在热带的海洋上。出现这种情况的原因，可能与热带海洋上空复杂的大气条件有关，在热带海洋地区，水汽丰富，大气多路径效应容易淹没反射信号；也可能与仪器性能有关，当直接无线电链路扫描大气最低层时，反射的信号是在掩星观测的最深处被捕获的，此时跟踪算法可能已经失效。

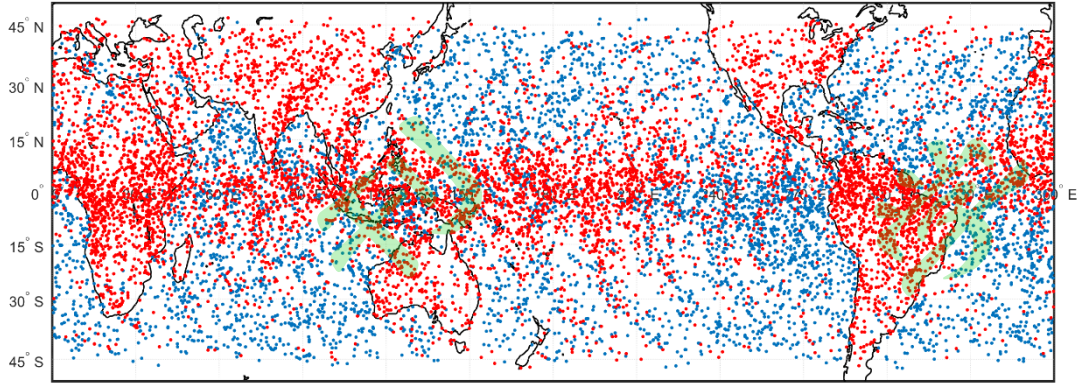


图 6 2020 年 1 月 1 日至 9 日的 13307 次 COSMIC-2 掩星事件的地理分布（蓝点代表有反射信号，红点代表无反射信号）

Fig. 6 The geographical distribution of 13,307 COSMIC-2 occultation events from January 1 to 9, 2020 (blue dots represent clear reflection signals, and red dots represent no-reflection signals)

4.2 反射信号与掩星折射率的偏差

本文把 COSMIC-2 的折射率廓线（atmPrf 文件）与 NCEP 预报值（avnPrf 文件）进行比较，检查两者的偏差与反射信号的关系。

我们将 2020 年 1 月 1 日至 9 日的 COSMIC-2 掩星事件分为 3 类：有清晰反射信号（5903 个）、有模糊反射信号（4519 个）和无反射信号（7404 个），然后与 NCEP 数据进行时间-空间匹配，得到 COSMIC-2 折射率-NCEP 折射率数据对，比较它们的平均偏差和标准差。我们选取 0.5~40km 高度范围，共分为 60 个高度层（0.5~5km 的间隔为 0.5km，5~40km 的间隔为 1km），在每个单独高度层 k 上计算折射率的偏差均值（ μ_k ）和标准差（ σ_k ）：

$$\mu_k = \frac{1}{M_k} \sum_{i=1}^{M_k} \frac{N_i^{\text{COSMIC-2}}(h_k) - N_i^{\text{NCEP}}(h_k)}{N_i^{\text{COSMIC-2}}(h_k)} \quad (5)$$

$$\sigma_k = \sqrt{\frac{1}{M_k} \sum_{i=1}^{M_k} \left(\frac{N_i^{\text{COSMIC-2}}(h_k) - N_i^{\text{NCEP}}(h_k)}{N_i^{\text{COSMIC-2}}(h_k)} - \mu_k \right)^2} \quad (6)$$

其中， M_k 为高度层 k 上的掩星事件数量， $N_i^{\text{COSMIC-2}}(h_k)$ 和 $N_i^{\text{NCEP}}(h_k)$ 分别为第 i 个 COSMIC-2/NCEP 的折射率廓线在第 k 层的值。计算结果如图 7 所示。

从图 7 可以看出，一般来说，在 5km 以下高度，三种类型的掩星事件折射率偏差均显著高于其他高度范围，且随着高度的降低掩星事件的数量快速减少，这对应于多路径传播效应的影响；8~30km 的折射率偏差均控制在 2% 以内，掩星反演的效果最好，此时多路径传播效应对 GPS 信号频率的影响微乎其微。

对于有清晰反射信号的掩星事件（图 7a），在 30km 以下高度的折射率偏差明显小于有模糊反射信号和没有反射信号的掩星事件的结果，特别是在 5km 以下高度，其折射率偏差控制在 3% 以内，与 NCEP 预测值拟合较好。

在 5km 以下高度，掩星反演的质量明显下降，由于反射信号包含低层大气和地表的信

息，可用于补充或改善 5km 以下的掩星反演结果，所以存在清晰反射信号的掩星折射率偏差相对较低。

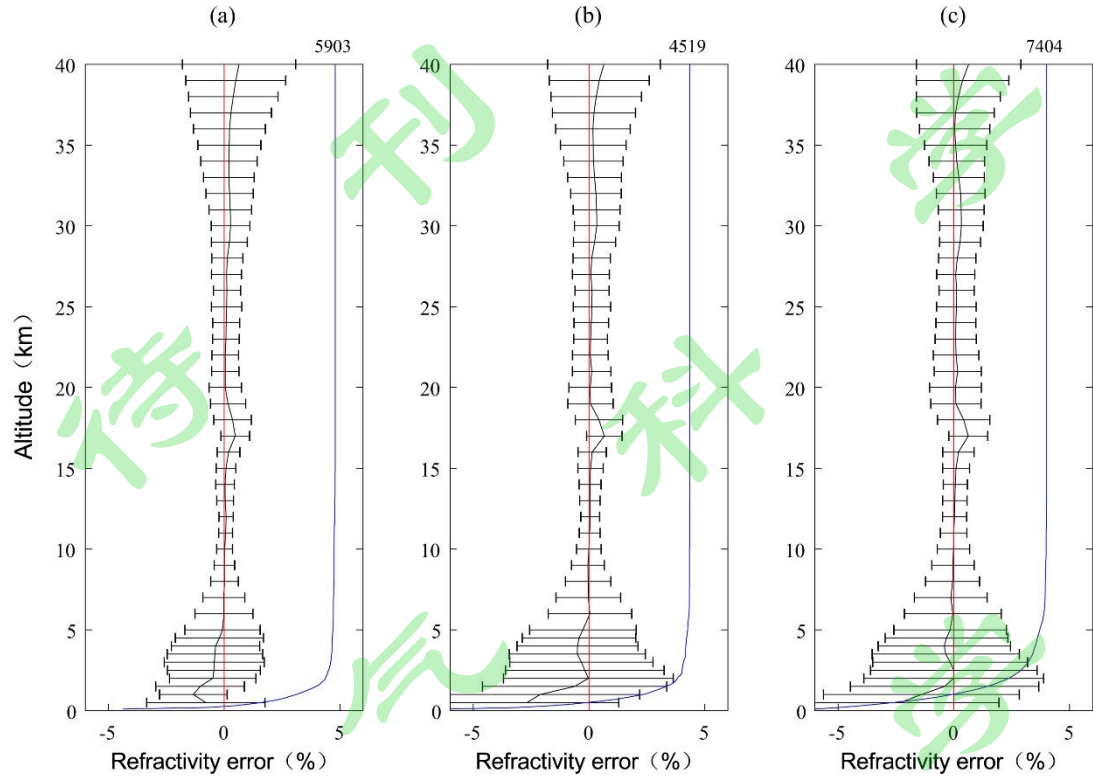


图 7 COSMIC-2 与 NCEP 比较的折射率廓线的误差均值和标准差（黑色曲线代表误差均值，黑色水平线段代表标准差，蓝色曲线代表掩星事件的数量）（a）有清晰反射信号的掩星事件（b）有模糊反射信号的掩星事件（c）没有反射信号的掩星事件

Fig. 7 The error mean and standard deviation of the refractive index profile compared between COSMIC-2 and NCEP (the black curve represents the mean error, the black horizontal line represents the standard deviation, and the blue curve represents the number of occultation events) (a) occultation events with clear reflection signals (b) occultation events with unclear reflection signals (c) occultation events with no-reflection signals

掩星事件的地理分布说明，反射信号的存在并没有干扰大气反演，而是证明仪器和大气条件足够好，换句话说，反射信号的存在是数据干净的证明，从而证明了大气探测的良好质量。通过分析反射信号与掩星折射率的偏差也发现，存在清晰反射信号的掩星折射率的偏差更小，特别是在 5km 以下高度。这些证据似乎表明：有反射信号的掩星事件数据质量更好，所包含的大气信息更丰富。

5 总结

GPS 掩星探测技术作为临边探测手段，具有自己独特的优点，为大气科学、空间物理学等的研究和应用提供了新的数据源。GPS 掩星信号中往往包含反射信号，如何从掩星信号中识别反射信号具有重要的意义。

本文提出了一种基于改进的 GoogLeNet 的反射信号分类模型，并应用于 COSMIC-2 掩

星探测数据，来识别掩星探测信号中的反射信号。首先选择 2020 年 1 月 1 日至 9 日的 COSMIC-2 掩星事件（conPhs 文件），通过质量控制得到下行掩星数据文件 17826 个，然后利用人工判读得到了明确有反射信号和没有反射信号的数据文件 13307 个；随机选择其中 70% 的数据作为训练集，15% 的数据作为验证集，剩余 15% 的数据作为测试集。结果表明，模型总体准确率达到了 96.4%，远优于传统的 SVM 方法。为了揭示反射信号的存在与掩星数据质量的关系，引入 NCEP 预报值（avnPrf 文件）与掩星资料进行比较，发现具有反射信号的掩星数据与 NCEP 数据的拟合更好，这进一步验证了前人的结论。

下一步，我们将通过选择合适的算法，进一步提升 Im-GNet 模型的性能，使之达到更好的反射信号识别效果；另一方面，由于包含反射信号的掩星数据含有更丰富的大气信息，因此，我们进一步研究如何利用反射信号所包含的信息来提高掩星探测数据质量。

参考文献（References）

- [1] Aparicio J M, Estel C, Hilda R. 2018. Information content in reflected signals during GPS Radio Occultation observations[J]. Atmospheric Measurement Techniques, 11(4):1883-1900.
- [2] Arne, Skov, Jensen, et al. 2004. Geometrical optics phase matching of radio occultation signals[J]. Radio Science, 39(3):1-8.
- [3] Beyerle G, and Hocke K. 2001. Observation and simulation of direct and reflected GPS signals in Radio Occultation Experiments[J]. Geophysical Research Letters, 28(9):1895-1898.
- [4] Beyerle G, Hocke K, Wickert J, et al. 2002. GPS radio occultations with CHAMP: A radio holographic analysis of GPS signal propagation in the troposphere and surface reflections[J]. Journal of Geophysical Research Atmospheres, 107(D24):1-14.
- [5] Boniface K, Aparicio J M, Cardellach E. 2011. Meteorological information in GPS-RO reflected signals[J]. Atmospheric Measurement Techniques, 4(7):1397-1407.
- [6] Cardellach, E. 2004. Carrier phase delay altimetry with GPS-reflection/occultation interferometry from low Earth orbiters[J]. Geophysical Research Letters, 31(10):377-393.
- [7] Cardellach E, Oliveras S, Rius A. 2008. Applications of the Reflected signals found in GNSS Radio Occultation Events [J]. Proceedings of GRAS SAF Workshop on Applications of GPS radio occultation measurements, 133-143.
- [8] Cardellach E, and Oliveras S. 2016. Assessment of a potential reflection flag product[R]. ROM SAF Report 23, http://www.romsaf.org/general-documents/rsr/rsr_23.pdf.
- [9] Gorbunov M E, Gurvich A S, Kornbluh L. 2000. Comparative analysis of radioholographic methods of processing radio occultation data[J]. Radio Science, 35(4):1025-1034.
- [10] Gorbunov M E. 2001. Radioholographic methods for processing radio occultation data in multipath regions[J]. Danish Meteorological Institute, 1-38.
- [11] Gorbunov M E. 2002. Radioholographic analysis of radio occultation data in multipath zones[J]. Radio

Science, 37(1):14-11.

- [12] Gorbunov M E, Lauritsen K B. 2004. Analysis of wave fields by Fourier integral operators and their application for radio occultations[J]. Radio Science, 39(4):1-15.
- [13] Gorbunov M E, Lauritsen K B, Rhodin A, et al. 2006. Radio holographic filtering, error estimation, and quality control of radio occultation data[J]. Journal of Geophysical Research: Atmospheres, 111(D10):2156-2202.
- [14] Gorbunov M E, Lauritsen K B, Leroy S S. 2016. Application of Wigner distribution function for analysis of radio occultations[J]. Radio Science, 45(6):1-11.
- [15] Gorbunov M E, Estel C, Lauritsen K B. 2018. Reflected ray retrieval from radio occultation data using radio holographic filtering of wave fields in ray space[J]. Atmospheric Measurement Techniques, 11(2):1181-1191.
- [16] Hocke K, Pavelyev A G, Yakovlev O I, et al. 1999. Radio occultation data analysis by radioholographic method[J]. Journal of Atmospheric and Solar-Terrestrial Physics, 61(15):1169-1177.
- [17] Ho S P, Anthes R A, Ao C O, et al. 2020. The COSMIC/FORMOSAT-3 Radio Occultation Mission after 12 years: Accomplishments, Remaining Challenges, and Potential Impacts of COSMIC-2[J]. Bulletin of the American Meteorological Society, 101(7):1107-1136.
- [18] Hu A, Wu S, Wang X, et al. 2018. Improvement of Reflection Detection Success Rate of GNSS RO Measurements Using Artificial Neural Network[J]. IEEE Transactions on Geoence & Remote Sensing, 56(2):1-10.
- [19] Jensen A S, Lohmann M S, Benzon H H, et al. 2003. Full Spectrum Inversion of radio occultation signals[J]. Radio Science, 38(3):6-1.
- [20] Kursinski E R, Hajj G A, Bertiger W I, et al. 1996. Initial Results of Radio Occultation Observations of Earth's Atmosphere Using the Global Positioning System (GPS)[J]. Science, 271(5252):1107-1109.
- [21] Kuo Y H, Wee T K, Sokolovskiy S, et al. 2004. Inversion and Error Estimation of GPS Radio Occultation Data[J]. Journal of the Meteorological Society of Japan, 82(1B):507-531.
- [22] Sokolovskiy S V. 2001. Modeling and inverting radio occultation signals in the moist troposphere[J]. Radio Science, 36(3):441-458.
- [23] Schreiner W S, Weiss J, Anthes R A, et al. 2020. COSMIC-2 Radio Occultation Constellation-First Results[J]. Geophysical Research Letters, 47(4):1-7.
- [24] 杨真真, 匡楠, 范露, 等. 2018. 基于卷积神经网络的图像分类算法综述[J]. 信号处理, 34(12):1474-1489.