

基于星载红外高光谱观测用机器学习算法反演大气温湿廓线

姚姝含 官莉

Atmospheric temperature and humidity profile retrievals using a machine learning algorithm based on satellite-based infrared hyperspectral observations

Yao Shuhan, Guan Li

在线阅读 View online: <https://doi.org/10.3788/IRLA20210707>

您可能感兴趣的其他文章

Articles you may be interested in

采用卷积神经网络的红外和可见光图像块匹配

Infrared-visible image patches matching via convolutional neural networks

红外与激光工程. 2021, 50(5): 20200364 <https://doi.org/10.3788/IRLA20200364>

基于深度谱卷积神经网络的高效视觉目标跟踪算法

Efficient visual target tracking algorithm based on deep spectral convolutional neural networks

红外与激光工程. 2018, 47(6): 626005 <https://doi.org/10.3788/IRLA201847.0626005>

基于卷积神经网络的反无人机系统图像识别方法

Image recognition method of anti UAV system based on convolutional neural network

红外与激光工程. 2020, 49(7): 20200154 <https://doi.org/10.3788/IRLA20200154>

流场测速中基于深度卷积神经网络的光学畸变校正技术

Aberration correction for flow velocity measurements using deep convolutional neural networks

红外与激光工程. 2020, 49(10): 20200267 <https://doi.org/10.3788/IRLA20200267>

基于3D卷积神经网络的PolSAR图像精细分类

Fine classification of polarimetric SAR images based on 3D convolutional neural network

红外与激光工程. 2018, 47(7): 703001 <https://doi.org/10.3788/IRLA201847.0703001>

基于神经网络的非线性大气修正实现红外目标辐射测量

Nonlinear atmospheric correction based on neural network for infrared target radiometry

红外与激光工程. 2020, 49(5): 20190413 <https://doi.org/10.3788/IRLA20190413>

基于星载红外高光谱观测用机器学习算法反演大气温湿度廓线

姚姝含, 官 莉

(南京信息工程大学 气象灾害教育部重点实验室/气象灾害预报预警
与评估协同创新中心, 江苏南京 210044)

摘 要: 星载红外高光谱垂直探测仪 GIIRS (Geostationary Interferometric Infrared Sounder) 能够实现大气温度和湿度参数高垂直分辨率的观测, 为数值天气预报提供精度更高的初始场。基于 GIIRS 观测辐射值采用 BP 神经网络 (Back Propagation Neural Network) 法和深度学习的卷积神经网络 (Convolutional Neural Networks, CNN) 法反演大气温度、湿度垂直廓线, 重点在于 CNN 法模型的构建与参数的优化, 得到反演精度最高的网络模型配置。将训练样本根据不同地表类型和是否有云的影响分为三种方案 (方案一: 不分类、方案二: 陆地/洋面分类、方案三: 晴空/有云分类), 分别进行建模、反演和检验。结果表明两种反演算法都有较好的反演精度, 相对而言 CNN 法在所有高度层上反演偏差、均方根误差和平均相对误差均较小, 反演精度更高。CNN 法温度反演在高层 10~200 hPa 改进较大, 三种分类方案改进的最大值分别为 1.15 K、1.06 K 和 1.02 K; 湿度反演在对流层低层 500~1 000 hPa 改进较大, 三种分类方案分别平均改进了 0.43 g/kg、0.41 g/kg 和 0.34 g/kg。BP 神经网络法方案三时 (即分晴空和云时) 温度和水汽混合比廓线反演精度最好; CNN 算法方案一时 (即不对样本数据进行任何分类) 反演精度最高。

关键词: GIIRS; 大气温湿度廓线; BP 神经网络; 卷积神经网络; 反演
中图分类号: P407.7 文献标志码: A DOI: 10.3788/IRLA20210707

Atmospheric temperature and humidity profile retrievals using a machine learning algorithm based on satellite-based infrared hyperspectral observations

Yao Shuhan, Guan Li

(Key Laboratory of Meteorological Disaster, Ministry of Education, Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: The satellite-based infrared hyperspectral Geostationary Interferometric Infrared Sounder (GIIRS) can achieve high vertical resolution observations of atmospheric temperature and humidity parameters, which provide a more accurate initial field for numerical weather forecasting. Based on GIIRS observation radiation, a back propagation (BP) neural network and deep learning convolutional neural networks (CNNs) are used to retrieve atmospheric temperature and humidity profiles, and the focus is on the construction of the CNN model and the optimization of parameters, thus obtaining the network model configuration with the highest retrieval accuracy. The training samples are divided into three schemes according to different surface types and the influence of whether there are clouds (scheme 1: no classification, scheme 2: land or ocean surface, scheme 3: clear or clouds)

收稿日期: 2021-09-26; 修订日期: 2021-11-07

基金项目: 国家自然科学基金 (41975028)

作者简介: 姚姝含, 女, 硕士生, 研究方向为卫星遥感资料的研究与应用。

导师简介: 官莉, 女, 教授, 博士生导师, 主要从事大气探测与大气遥感方面的研究。

and modelling, retrieving and testing. The results show that the two retrieval algorithms both have good retrieval precision. Relatively speaking, the CNN method has a smaller retrieval bias, root-mean-square error and mean relative error at all altitudes, and the retrieval precision is higher. The temperature retrieval of the CNN method is greatly improved in the high level at 10-200 hPa, and the maximum values of the three classification schemes are 1.15 K, 1.06 K, and 1.02 K, respectively, and the humidity retrieval of the CNN method also shows improvement in the lower troposphere at 500-1 000 hPa, and the averages of the three classification schemes are 0.43 g/kg, 0.41 g/kg, and 0.34 g/kg, respectively. The third scheme (clear or clouds) of the BP neural network method has the best retrieval precision of temperature and water vapour mixing ratio profiles, and the first scheme (no classification of sample data) of the CNN algorithm has the most accurate retrieval results.

Key words: GIIRS (Geostationary Interferometric Infrared Sounder); atmospheric temperature and humidity profile; BP neural network; CNN (Convolutional Neural Networks); retrieval

0 引言

大气温度和湿度廓线是描述大气状态不可缺少的参数,也是气象预报和气候预测研究及业务的主要参考因素,获取准确的大气温湿度廓线信息对数值天气预报和气候变化研究具有重要的意义^[1],并且对于短期的天气预警也有一定改善。获取大气温湿度廓线信息的传统方法是通过气象站点的探空气球观测所得,虽然该方法可以获得较精确的大气温湿度廓线信息,但是由于地理条件等各方面的因素,探空观测仅限于站点观测,观测得到的温湿度廓线信息的空间分辨率和时间分辨率都比较低,因此已经不能满足业务及现代气象发展的需求。但是气象卫星能够连续地、均匀地观测全球大气状况,弥补沙漠、广阔海洋、高原和两极地区探空资料的不足,获得全球大气温湿度分布资料。

随着气象卫星的不断发展,红外高光谱垂直探测器开辟了气象卫星探测大气的新时代,由于可以探测到更窄光谱分辨率的大气辐射特征,其探测的信息量是一般气象卫星的上百倍,能够得到更加精细的大气结构,提高了探测大气垂直结构的能力和精度,尤其是温度和湿度垂直廓线^[2]。为实现红外大气探测技术在数值预报中的应用,国内外都开始了红外高光谱垂直探测仪的研制、资料处理等应用方法的研究。

星载红外高光谱垂直探测仪探测到的信息并不是直接的大气温度、湿度等,而是来自地球大气系统的红外辐射,为了使大气遥感的量化水平满足数值预报等应用越来越高的要求,国内外的科学家不断研究遥感反演新方法,不断提高遥感反演精度,期望能够达到数值模式直接应用大气温湿等参数的精度要

求^[3]。基于星载红外高光谱垂直探测仪资料反演大气温湿度廓线常采用的传统方法主要有统计回归法和物理反演法。例如:Smith等^[4]使用一种依赖于正交经验函数(EOF)的双回归反演算法基于CrIS(Cross-track Infrared Sounder)观测数据,进行大气温湿度廓线反演,其温度廓线反演均方根误差在2 K以内,湿度廓线反演均方根误差在20%以内。官元红等^[5]利用IASI(Infrared Atmospheric Sounding Interferometer)高光谱传感器温度和水汽探测通道资料,使用一维变分物理方法,建立大气温度和湿度廓线变分反演系统,并选取北京、青岛、沈阳三个地区展开了反演试验,将其反演结果与探空数据进行对比,温湿度廓线反演均方根误差分别为0.89 K、0.02 g/kg。Zhu等^[6]用HIRAS(Hyperspectral Infrared Atmospheric Sounder)基于一维变分算法反演大气温湿度廓线,用观测亮温和模型模拟亮温之间的误差来校正观测数据,并用校正后的观测数据进行反演,温度廓线反演均方根误差小于1.5 K,湿度廓线小于10%。近年来,随着人工智能科学技术的发展,在气象大数据的背景下,机器学习算法逐渐被引入大气科学领域。机器学习算法具有非线性映射能力,善于从输入和输出信号中寻找规律,不需要精确的数学模型,相比传统的统计回归方法能更好的解决非线性问题,且不需要对非线性映射本身做过多的分析。例如浅度学习的人工神经网络算法被广泛应用于卫星遥感反演大气温湿度廓线或地表参数^[7-8]。神经网络反演算法是对卫星观测和大气参数建立一个非线性的统计模型,并且该网络模型具有极强的非线性问题处理能力和较好的容错能力,使得不仅能够提高反演精度又可以节省计算时

间。与辐射传输模式不同,神经网络算法不需要依赖于物理过程的知识,就可以获得卫星数据输入与大气信息的最佳模型^[9]。官莉、刘旻^[10-11]基于 AIRS (Atmospheric Infrared Sounder) 资料利用神经网络法反演大气温湿度廓线,结果表明较特征向量统计法该方法提高了反演精度,并且神经网络反演法更能揭示真实大气湿度廓线的细微结构。Huang^[12]等结合神经网络法和改进的一维变分算法,使用 GIIRS 数据反演大气温度廓线,以怀俄明大学发布的探测数据作为真值验证反演精度,温度反演均方根误差小于 2 K 且其精度优于 GIIRS 温度产品。Adam B. Milstein 和 William J. Blackwell^[13]提出了一种改进的神经网络算法,用于从微波和红外高光谱观测中反演大气温湿度廓线,验证结果表明该算法在反演误差和垂直分辨率方面都有显著提高。随着卫星遥感和机器学习的不断发展,深度学习的机器学习算法也逐渐被应用于卫星遥感反演,例如 David Malmgren-Hansen^[14]等利用 IASI 观测数据,用卷积神经网络算法反演大气温湿度廓线,并与线性回归方法进行比较,发现 CNN 法反演精度更高,表明深度学习算法是一种更适用于大气廓线反演的机器学习算法。国内外深度机器学习算法在大气温湿度廓线反演中的应用研究成果都较少。

目前关于星载高光谱红外探测器大气温湿度廓线反演的研究主要集中于极轨卫星平台上,如上述 CrIS、IASI、AIRS 等,而中国第二代静止气象卫星 FY-4 A 于 2016 年 12 月 11 日成功发射,首次在静止卫星上装载了干涉式大气红外高光谱垂直探测器 GIIRS (Geostationary Interferometric Infrared Sounder),其主要目的是实现大气温湿参数的垂直探测,提高气象观测的垂直分辨率^[12]。同时结构框架复杂的深度机器学习算法(即包含多级非线性变换的层级机器学习方法)在大气温湿度廓线反演应用中很少见文献报导。如深度学习的 CNN 卷积神经网络主要通过结构重组和减少权值将特征提取融合进多层感知器,将特征表示和回归预测统一起来,在特征提取和建模上都具有相较于浅层模型显然的优势。因此,文中针对 GIIRS 观测,拟同时采用 BP 神经网络方法以及深度学习的 CNN 方法对大气温度、湿度垂直廓线的反演进行模拟研究。考虑到 CNN 网络结构框架复杂且网络参数较多,文中侧重于 CNN 网络结构和参数的优

化,以提高网络的适应性和反演精度。

1 数据与预处理

1.1 训练样本

文中采用由美国威斯康星大学开发的用于高光谱和多光谱反演的全球大气晴空训练样本 CIMSS (Cooperative Institute for Meteorological Satellite Studies),该样本共由全球范围内的 15 704 条大气温度、湿度和臭氧廓线等组成,包括了 NOAA-88、TIGR-3、Radiosondes、Ozonesondes 以及 ECMWF 等具有代表性的五个来源的资料。每条廓线均经过了严格的质量控制和检验,并从 1 100~0.005 hPa 被分成 101 层,具有较高的垂直分辨率,此外,每条廓线都有相对应的地表发射率和地表温度,能够更精确地从卫星观测辐射值反演大气温湿度廓线^[15]。该样本不仅具有代表性且保证了样本数量,已被多次应用于红外高光谱大气参数反演^[16]。由于 GIIRS 探测器主要观测范围为中国区域,而南北两极地区地形及其气候等条件较为复杂,且不覆盖中国区域,所以文中只选择南北纬 60°范围内的 12 528 条大气廓线作为训练样本库。

GIIRS 观测共覆盖两个谱段,分别为 700~1 130 cm^{-1} 的长波波段和 1 650~2 250 cm^{-1} 的短中波波段,两个波段的最高光谱分辨率为 0.625 cm^{-1} ,其中长波波段 689 个通道,中波波段 961 个通道,共有 1 650 个观测通道。

快速辐射传输模式使用的是由欧洲中期天气预报中心开发的 RTTOV (Radiative Transfer for Tiroson Operational Vertical Soinder),该模式可以快速且精准地模拟出各类卫星仪器在给定大气状态参数条件下的观测亮温值^[17]。

将 CIMSS 廓线样本集的 101 层温度、湿度、臭氧廓线数据作为快速辐射传输模式的输入,通过 RTTOV 模式模拟出 GIIRS 1 650 个通道的观测亮温值。为确保模拟的真实性,将对模拟出的亮温值增加一些观测噪音,噪音根据高斯分布随机产生,平均值取零,标准差为 GIIRS 的 NeDT (Noise Equivalent Delta Temperature, 等同于 300 K 亮温值时的噪声值)。

1.2 反演通道选择

训练样本库由 12 528 对模拟的 GIIRS 通道亮温值和相应的 101 层温度和湿度廓线对组成,包括训练样本和测试样本。考虑到红外高光谱仪器通道间存

在很高的相关性,以及用如此大的输入量进行网络训练,不光训练时间长,而且实际反演时计算时间花费也高,所以文中先对 GIIRS 的 1650 个通道进行通道选择,从中选出适合反演大气温度和湿度廓线的最佳通道组合。

参考 Gambacorta 等公布的 CrIS 的 399 个温度、湿度反演通道合集^[8],先选出其中与 GIIRS 通道中心波长完全相同的 64 个温度通道、10 个水汽通道、53 个臭氧通道、26 个 HNO₃ 通道、17 个 N₂O 通道和 15 个 CO 通道。由于 CrIS 仪器设置的水汽吸收带在 1210~1750 cm⁻¹,而 GIIRS 的水汽吸收带使用的是 1650~2155 cm⁻¹,因此需在该水汽吸收波段中选出适用于 GIIRS 大气温湿度廓线反演的通道。具体步骤:计算 1650~2155 cm⁻¹内每个通道的权重函数(共 809 个通道);在 1100~150 hPa 高度范围内寻找权重函数峰值分别分布在每个高度层(40 个高度层)上的通道;从各高度层中选择权重函数值最大的通道作为水汽通道。这样保证在每个垂直分层上都有一个权重函数峰值最大的通道,依据该方法在 GIIRS 水汽吸收带选出了 40 个水汽通道,最终选取 225 个通道用于大气温湿度廓线反演。通道分布如图 1 所示,其中包括 64 个温度通道(绿色)、50 个水汽通道(红色)、53 个臭氧通道(蓝色)、26 个 HNO₃ 通道(紫色)、17 个 N₂O 通道(黄色)和 15 个 CO 通道(蓝绿色)。

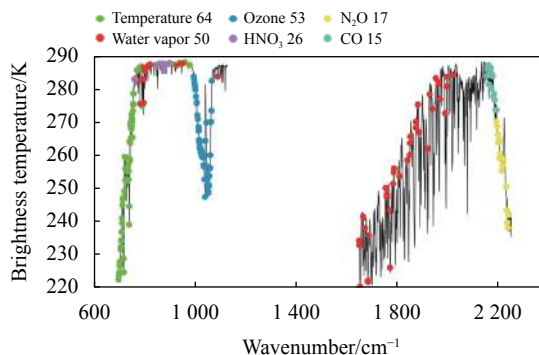


图 1 选择的 GIIRS 通道

Fig.1 Selected GIIRS channels

2 大气温湿度廓线反演

2.1 算法流程

采用 BP 神经网络反演法和 CNN 反演法实现 GIIRS 大气温湿度廓线反演,两种方法实现流程如图 2 所示。先将训练样本集中大气温度、湿度和臭氧廓线数据作为 RTTOV 辐射传输模式输入,模拟得到 GIIRS 辐射通道亮温值,从中选择出 1.2 节中 225 个反演通道的亮温数据,分别组成温度廓线反演样本对和湿度廓线反演样本对。之后将反演样本对均匀间隔选取 1000 个样本对作为独立的测试样本,剩余的 11528 个样本对作为训练样本,并用 BP 神经网络法和 CNN 法分别建立反演模型,通过对两种模型框架及参数不

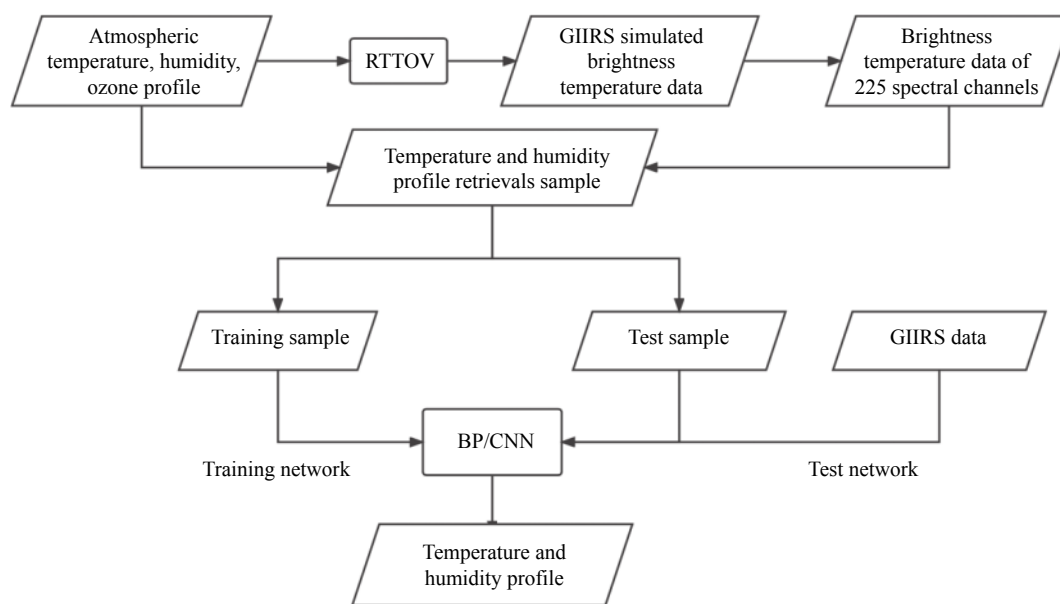


图 2 方法实现流程图

Fig.2 Method implementation flowchart

断的优化和调整,得到最佳的反演模型。最终用 1000 个独立测试样本的亮温数据作为反演模型的输入,得到反演的大气温湿度廓线,将反演结果与测试样本廓线“真值”进行反演结果精度测试,并对两种反演算法进行比较和分析。

2.2 BP 神经网络反演法

BP (Back propagation Neural Network) 神经网络是目前人工神经网络算法中应用最广泛的模型,能够实现任意精度的连续函数映射,有效用于复杂的非线性函数的逼近^[10]。文中采用的是一个三层的前馈网络,包括输入层、隐含层和输出层。选择的 225 个通道的模拟亮温值作为该 BP 神经网络的输入层,对应的 101 层大气温度、湿度廓线分别作为输出层,即输入层有 225 个节点,输出层有 101 个节点。隐含层设置为一层,网络性能的好坏受隐含层节点数设置的影响,如果隐含层节点数太少,会使得信号资料不够,则必定会影响网络的效果;而隐含层节点数太多,则会使得训练花费更多的时间,降低业务效率。基于以前的研究^[19-21],目前应用较广泛的关于隐含层节点数选取的方法有三种,分别为公式(1)~(3),经过试验及比较,对于隐含层节点数的设置主要参考了公式(3),即隐含层节点为 108 个。

$$h = \log_2 T \quad (1)$$

$$h = \sqrt{mn} \quad (2)$$

$$h = \sqrt{(0.43mn + 0.12m^2 + 2.54n + 0.77m + 0.35)} + 0.51 \quad (3)$$

式中: T 为训练样本的个数; h 为隐含层的节点数; m 为输出层的节点数; n 为输入层的节点数。

使用 Newff 函数创建神经网络,带动量的梯度下降法 traingdx 作为网络的训练函数,双曲正切 S 型传递函数 tansig 作为网络的激活函数。神经网络训练参数如表 1 所示。

表 1 BP 神经网络训练参数

Tab.1 Training parameters of BP neural network

Parameter	Set value	Attributes
Net.trainParam.epochs	10000	Training times
Net.trainParam.goal	0	Training goal
Net.trainParam.lr	0.01	Learning rate
Net.trainParam.mc	0.95	Momentum factor
Net.trainParam.show	25	Number of intervals displayed
Net.trainParam.min_grad	1×10^{-6}	Minimum performance gradient

2.3 CNN 反演法

2.3.1 模型结构

卷积神经网络 (Convolutional Neural Networks, CNN) 在图像识别与分类、语义分割、智能驾驶等领域已有优异的表现^[22-23],它是一种包含了卷积计算并且有深度结构的前馈神经网络,是深度学习算法代表之一。它也是一种特殊的深层神经网络模型,其特殊性主要体现在两方面:一方面是它的神经元之间的连接是非全连接的;另一方面是在同一层中某些神经元间的连接权重是共享的。它的非全连接和权值共享的网络结构降低了网络模型的复杂度,减少了权值的数量。

CNN 的结构主要包括输入层、隐含层和输出层。其中隐含层主要包括卷积层、池化层、全连接层这三种常见的网络层,隐含层是实现特征提取的关键技术。卷积层的主要功能就是对输入进行特征提取,它的内部包含多个卷积核,卷积层内的每个神经元都与前一层中位置接近的区域的多个神经元相连,区域的大小取决于卷积核的大小。卷积核在进行卷积工作时相当于滤波器按设定的步长在输入图片上进行滑动操作,输入图片与卷积核进行卷积计算,并产生与卷积核个数相同的特征映射图。卷积运算如下:

$$y^{(l+1)} = k^{(l)} x^{(l)} + b^{(l)} \quad (4)$$

式中: $x^{(l)}$ 为上层的输出; $k^{(l)}$ 为 l 层中的某个卷积核; $b^{(l)}$ 为偏置。则该层的输出为:

$$x^{(l+1)} = f(x^{(l)}) \quad (5)$$

式中: $f(*)$ 为激活函数。

池化层相当于一个降采样的过程。输入的图片在经过卷积层内不同卷积核的卷积运算后,输出的特征图会被传递到池化层进行特征选择和信息过滤,池化层通过对不同位置的特征作聚合统计,得到维度较低的统计特征^[24],目前常用的池化操作有最大值池化和平均值池化。全连接层从网络结构上讲等同于传统神经网络中的隐含层,位于卷积神经网络隐含层的最后部分,与卷积层和池化层的局部连接方式不同的是全连接层的每一个节点都需与上一层的所有节点相连,从而将前面几层提取到的局部特征进行非线性组合综合得到全局特征。卷积神经网络输出层的上一层通常都是全连接层,其工作原理和结构与传统的

前馈神经网络的输出层相同。

2.3.2 模型参数设置

采用的模型包含一个输入层, 四个卷积层、两个池化、一个全连接层和一个回归输出层, 卷积层和池化层交替设置构成一个多层神经网络, 框架结构如图 3 所示, 图中卷积神经网络的输入层为每个样本的

225 个通道亮温值, 可以将其视为宽为 1 的一维图像, 每个像素点为每个通道的亮温, 所以输入层大小为 225×1 , 输出层为 1×101 的大气温度、湿度标签, 用一维卷积核进行卷积操作, 图中深色部分为卷积核大小。卷积层主要提取样本数据特征, 将会输出若干个特征图 (feature maps), 是卷积神经网络中最核心的部分。

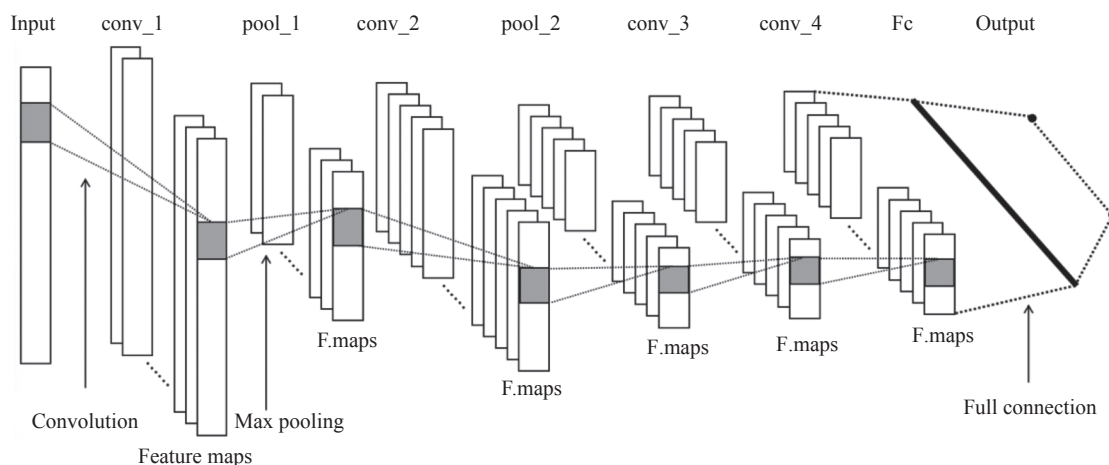


图 3 CNN 模型框架结构图

Fig.3 CNN model frame structure

由于深度学习的 CNN 网络框架复杂、模型参数众多, 实际应用中, 为了得到最优的模型参数, 需对其进行反复调试。以温度反演模型为例, 计算不同参数设置时网络的训练误差。参考指标为验证均方根误差 RMSE、温度反演 RMSE 和网络训练时间, RMSE 越小、训练时间越短则网络性能越好。其中: 验证 RMSE 为网络提供的检验参数, 温度反演 RMSE 为用训练库中均匀间隔选取的 1000 对独立测试样本统计的反演误差。以第一卷积层 (conv_1) 为例, 测试选取不同卷积核大小时的参考指标如表 2 所示, 其中第一列为卷积核大小, 第二列为验证 RMSE, 第三列为独立检验样本的温度反演 RMSE, 第四列为网络训练时长, 发现卷积核越大则网络训练时间越长, 而验证 RMSE 和温度反演 RMSE 则越小, 综合考虑参考指标和输入信号的大小, 最终第一卷积层的卷积核大小设置为 5×1 的一维卷积核。测试第一卷积层最优输出特征图 (feature maps) 个数的参考指标如表 3 所示, 同表 2, 发现输出特征图数越多则验证和温度反演 RMSE 越小、网络训练时间越久, 但是输出特征图数从 50 到 60 时, 验证 RMSE 和温度反演 RMSE 不再减小, 训练时间却明显加长, 所以最终选择第一卷积层

输出特征图数为 50。

其他网络参数最优选择与上述方法类似, 不再赘述。每一个卷积层都选用同样大小的滤波器 (即卷积核大小 5×1 的一维卷积核), 卷积核移动的步幅为 1, 池化层采用大小为 2×1 、步幅为 2 的平均值池化方式。卷积层中每个神经元的输入通过激活函数映射到输出端, 增加网络模型的非线性, 常用的激活函数有 sigmoid、tanh 和 ReLU 函数, 其中 sigmoid 和 tanh 函数由于都存在饱和的问题, 容易造成梯度消失, 而 ReLU 函数为不饱和的非线性函数, 可以避免梯度消失、梯度爆炸等问题, 同时在网络训练时具有更快的收敛速度, 因此激活函数选为 ReLU 函数。为了加快

表 2 第一卷积层不同卷积核大小时网络训练参考指标
Tab.2 Network training reference indicators for different convolution kernel sizes of conv_1

Convolution kernel size	Validation RMSE	Temperature retrieval RMSE/K	Training time
3×1	0.8521	3.4625	12'48"
4×1	0.8329	3.4217	13'14"
5×1	0.8298	3.4137	14'20"
6×1	0.8274	3.3871	15'00"

表 3 第一卷积层不同输出特征图数时网络训练参考指标

Tab.3 Network training reference indicators for different output feature maps of conv_1

Number of output feature maps	Validation RMSE	Temperature retrieval RMSE/K	Training time
20	0.791 6	3.288 8	16'14"
30	0.781 6	3.242 7	22'47"
40	0.773 0	3.224 5	27'32"
50	0.759 5	3.169 8	27'55"
60	0.761 7	3.175 6	35'80"

网络的训练速度,网络训练优化器选 Adam^[14],与传统的优化器相比,Adam 具有更高的计算效率以及较低的内存需求。文中 CNN 网络模型完整构架及参数如表 4 所示。

表 4 CNN 模型完整构架

Tab.4 Complete structure of CNN model

Name	Type	Output	Parameter
imageinput	Image input	225×1×1	-
conv_1	Convolution	225×1×50	5×1×1×50
batchnorm_1	Batch normalization	225×1×50	1×1×50
relu_1	ReLU	225×1×50	-
avgpool_1	Average pooling	112×1×50	-
conv_2	Convolution	112×1×100	5×1×50×100
batchnorm_2	Batch normalization	112×1×100	1×1×100
relu_2	ReLU	112×1×100	-
avgpool_2	Average pooling	56×1×100	-
conv_3	Convolution	56×1×100	5×1×100×100
batchnorm_3	Batch normalization	56×1×100	1×1×100
relu_3	ReLU	56×1×100	-
conv_4	Convolution	56×1×100	5×1×100×100
batchnorm_4	Batch normalization	56×1×100	1×1×100
relu_4	ReLU	56×1×100	-
dropout	Dropout	56×1×100	-
fc	Fully connected	1×1×101	101×5600
regressionoutput	Regression output	-	-

3 反演结果检验

将上述均匀间隔选取的 1000 个独立的测试样本,分别用 BP 神经网络法和 CNN 法建立的模型进行大气温度、湿度垂直廓线反演,并统计反演误差。精度评价指标有平均误差 ME、均方根误差 RMSE 和平均相对误差 MRE,定义分别为:

$$ME = \frac{1}{N_s} \sum_{i=1}^{N_s} (y_i - x_i) \quad (6)$$

$$RMSE = \sqrt{\frac{1}{N_s} \sum_{i=1}^{N_s} (x_i - y_i)^2} \quad (7)$$

$$MRE = \frac{\frac{1}{N_s} \sum_{i=1}^{N_s} |(x_i - y_i)|}{\frac{1}{N_s} \sum_{i=1}^{N_s} x_i} \times 100\% \quad (8)$$

式中: x_i 为检验样本廓线; y_i 为 GIIRS 反演的大气廓线; N_s 为检验样本个数。

针对不同地表类型和是否有云的影响对 BP 法和 CNN 法分别建立了三种网络模型:第一方案为对样本廓线不做任何方式的分类;第二方案将样本分为陆地和海洋,分别进行建模和反演;第三方案参考 ZH-A10 方法^[25]将样本分为有云和晴空分别建模和反演。

图 4 显示了整层大气 (0.005~1 100 hPa) 的温度反演值和“真值”的散点分布图,其中图 4(a) 不分类、图 4(b) 陆地/洋面分类、图 4(c) 晴空/云分类为 BP 神经网络法得到的散点图;图 4(d)~(f) 使用的是 CNN 算法,横坐标为测试样本温度值,纵坐标为反演温度值。从图中可以看出,两种方法反演值和“真值”有较高的一致性(相关系数高达 0.99),均分布在直线 $y=x$ (图中红线) 的两侧。BP 神经网络法三种分类方案的相关系数 R 均为 0.989,平均误差 ME 分别为 0.01 K、0.006 K、0.008 K,均方根误差 $RMSE$ 分别为 3.96 K、4.03 K、3.92 K,平均相对误差 MRE 分别为 0.95%、0.98%、0.94%;CNN 法三种分类方案的 R 均为 0.99, ME 分别为 -0.02 K、-0.0008 K、0.01 K, $RMSE$ 分别为 2.90 K、3.03 K、3.07 K, MRE 分别为 0.65%、0.69%、0.69%。从 R 、 ME 、 $RMSE$ 和 MRE 都可以看出,CNN 法优于 BP 神经网络法。图 5 与图 4 相似,显示了水汽混合比的散点图。同样,CNN 法反演的 R 、 ME 、 $RMSE$ 和 MRE 均优于 BP 神经网络法。至于不同分类方案,BP 神经网络法用方案三(即将样本分为有云和晴空)时得到的反演结果较其他两种方案稍有提高,均方根误差较小;而 CNN 法不对样本数据进行任何分类时其反演均方根误差最小。

图 6 和图 7 分别显示了检验样本统计的温度反演平均误差、均方根误差和平均相对误差垂直廓线,

其中 (a)、(b)、(c) 分别为三种不同分类方案, 虚线代表平均误差, 黑色代表均方根误差, 图 7 则为平均相对误差。从图中可以看出, 两种算法的平均误差在所

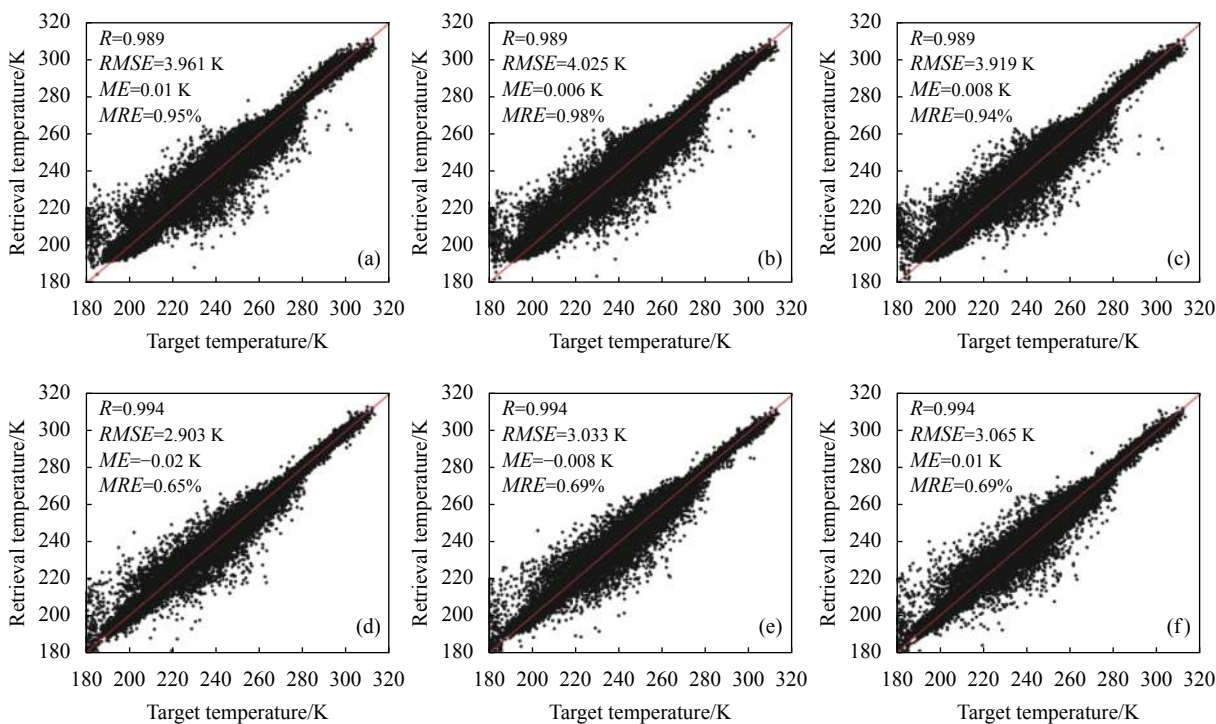


图 4 温度反演散点图。(a)-(c) BP 神经网络三种分类方案; (d)-(f) CNN 三种分类方案

Fig.4 Temperature scatter diagram of retrieval. (a)-(c) Three classification schemes of BP neural network; (d)-(f) Three classification schemes of CNN

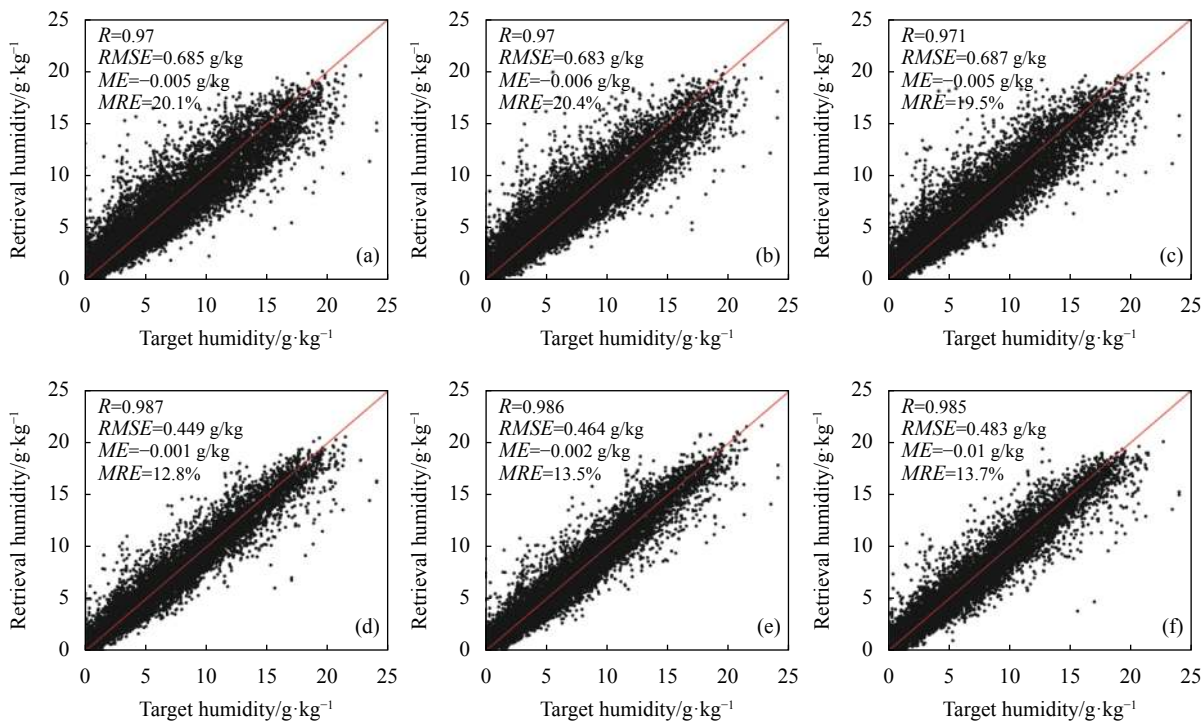


图 5 同图 4 水汽混合比

Fig.5 Water vapor mixing ratio same as Fig.4

有高度上相当,均较小;相较于BP神经网络法,CNN法就均方根误差RMSE和平均相对误差MRE而言均有明显的改进,在高层10~200 hPa改进较大,三种分

类方案RMSE改进的最大值分别为1.15 K、1.06 K和1.02 K,MRE改进的最大值分别为0.45%、0.42%和0.39%,对流层低层600~1 000 hPa均方根误差减小较

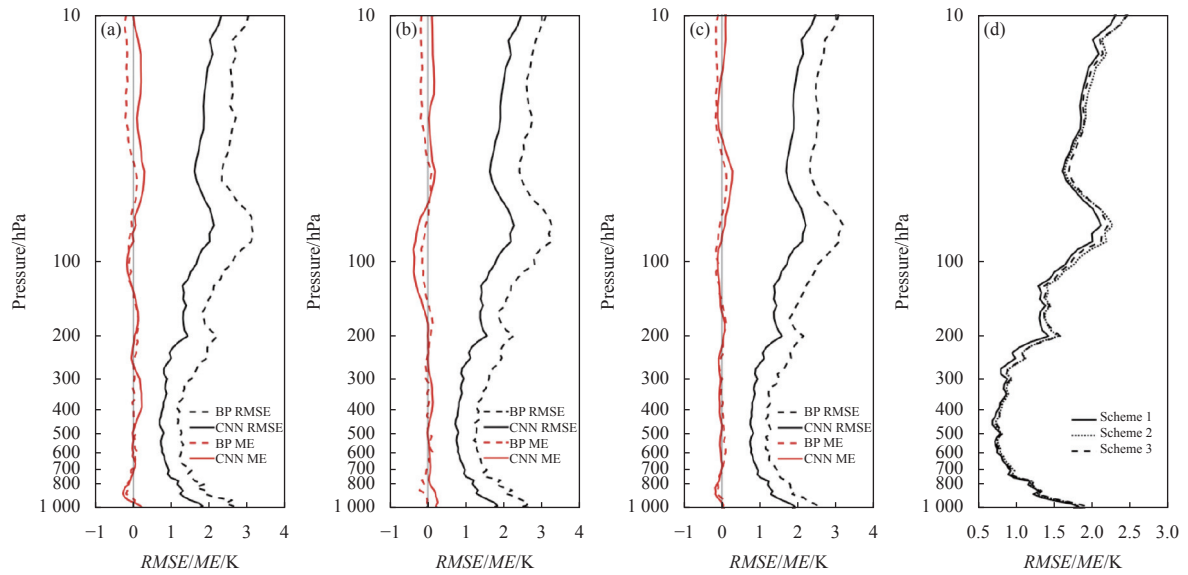


图 6 温度反演误差廓线。(a)~(c) 三种分类方案,红色代表偏差,黑色代表均方根误差,虚线代表BP神经网络法,实线代表CNN法;(d) CNN 三种分类方案均方根误差廓线,实线方案一,点画线方案二,虚线方案三

Fig.6 Error profile of retrieval for temperature. (a)-(c) Three classification schemes, red is the bias, black is the root mean square error, dotted line is the BP neural network method, and the solid line is the CNN method; (d) Root mean square error profile of the three classification schemes of CNN, the solid line is the first scheme, the dotted line is the second scheme, and the dashed line is the third scheme

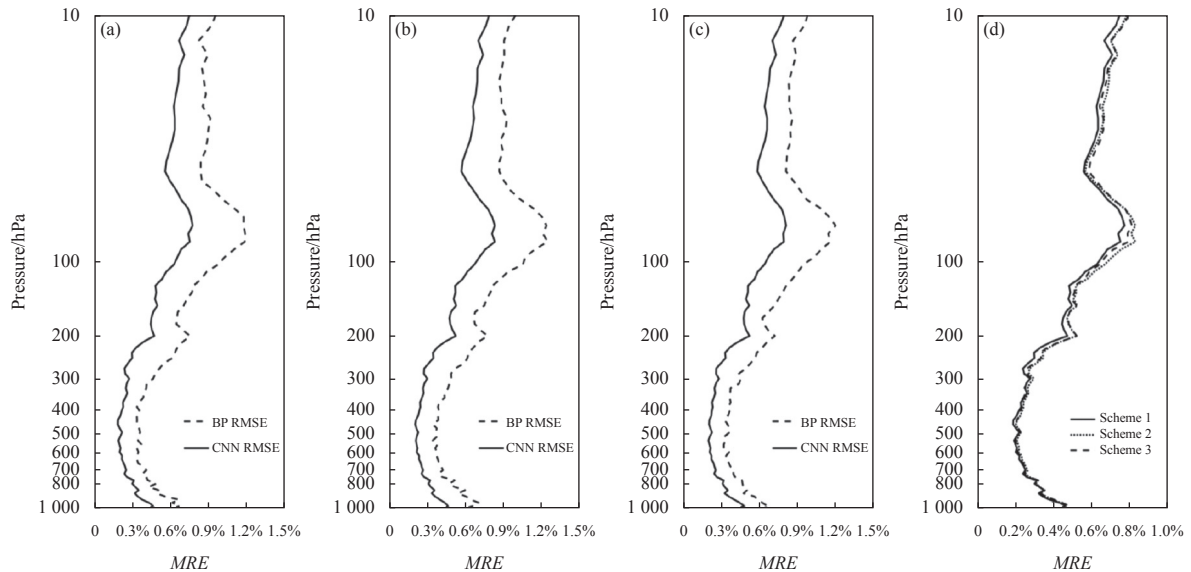


图 7 温度反演平均相对误差廓线。(a)~(c) 三种分类方案,虚线代表BP神经网络法,实线代表CNN法;(d) CNN 三种分类方案平均相对误差廓线,实线方案一,点画线方案二,虚线方案三

Fig.7 Mean relative error profile of retrieval for temperature. (a)-(c) Three classification schemes, dotted line is the BP neural network method, and the solid line is the CNN method; (d) Mean relative error profile of the three classification schemes of CNN, the solid line is the first scheme, the dotted line is the second scheme, and the dashed line is the third scheme

小, *RMSE* 减小了约 0.6 K, *MRE* 减小了约 0.17%。为进一步分析不同分类方案的差别, 两张图中的 (d) 分别给出了 CNN 法三种分类方案时温度反演均方根误差和平均相对误差垂直廓线(即 (a)~(c) 图中的黑色实线), 其中实线代表方案一, 点画线代表方案二, 虚线代表方案三。可以看出, 方案一反演结果较其他两种方案在所有高度层均方根误差和平均相对误差都偏小, 而分类方案二和三则在不同高度上各有优劣。

图 8 和图 9 分别给出了水汽混合比反演的平均

误差、均方根误差和平均相对误差垂直廓线, 同样, CNN 法比 BP 神经网络法反演的湿度廓线误差小, *RMSE* 在对流层低层 500~1000 hPa 改进较大, 三种分类方案分别平均改进了 0.43 g/kg、0.41 g/kg 和 0.34 g/kg, *MRE* 在各层改进均较大, 三种分类方案分别在各层平均改进了 7.06%、7.41% 和 5.56%。两张图中的 (d) 同样分别显示了不同分类方案水汽混合比均方根误差和平均相对误差垂直廓线, 对 CNN 法而言, 在所有高度层方案一反演精度均高于其他两种分类方案, 尤

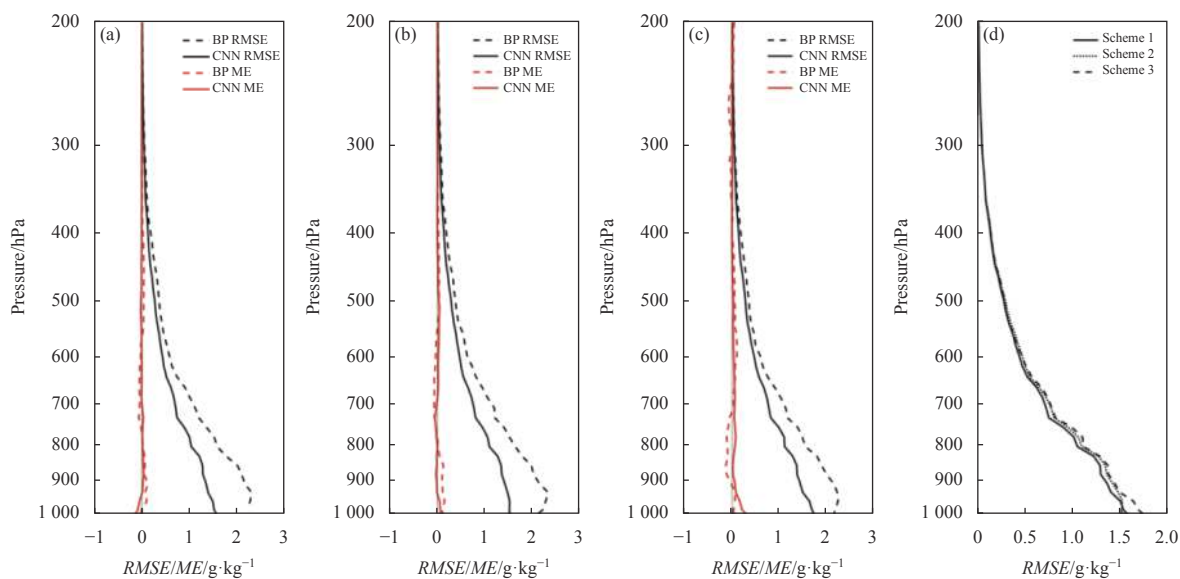


图 8 同图 6 水汽混合比

Fig.8 Water vapor mixing ratio same as Fig.6

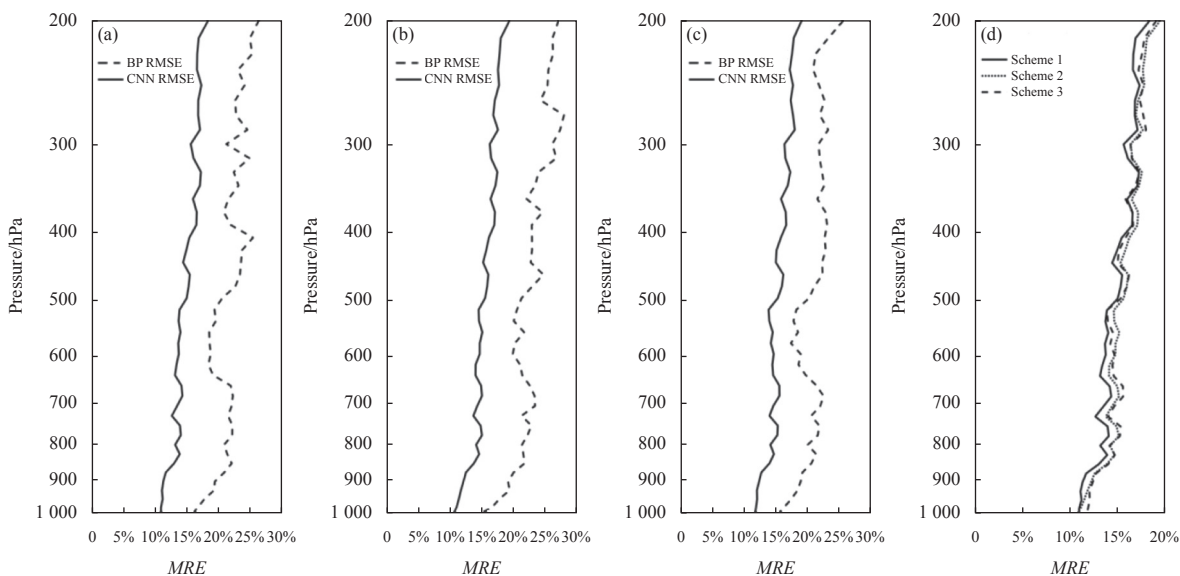


图 9 同图 7 水汽混合比

Fig.9 Water vapor mixing ratio same as Fig.7

其低层 500~1 000 hPa 改进最大。

4 结 论

基于星载红外高光谱探测仪 GIIRS 模拟观测资料,分别用浅层的 BP 神经网络法和深度 CNN 机器学习算法反演大气温度、湿度廓线,测试最佳的网络框架结构和网络参数,并比较了两种算法的反演精度。得到如下结论:

(1) 就 GIIRS 大气温湿度廓线反演而言, BP 神经网络法和 CNN 法均取得较高的反演精度;机器学习算法相较于物理反演算法节省了计算时间,应用前景较好。

(2) 深度学习的 CNN 模型框架结构复杂且网络参数较多,为使反演精度最高,经过反复测试和优化,构建了包含一个输入层、四个卷积层,两个池化层、一个全连接层和一个回归输出层的大气温度、湿度廓线反演框架。其中每个卷积层卷积核大小选为 5×1 ,池化采用大小为 2×1 、步幅为 2 的平均池化方式,激活函数为 ReLU,网络优化器为 Adam。

(3) 不论是温度还是水汽混合比反演,深度学习的 CNN 法均优于传统的 BP 神经网络法,其反演精度在所有大气高度层都有一定程度的提高。10~1 000 hPa 温度反演 RMSE 三种分类方案分别平均改进了 0.679 K、0.698 K 和 0.603 K, MRE 分别平均改进了 0.23%、0.23% 和 0.20%;湿度反演在 200~1 000 hPa 三种分类方案 RMSE 分别平均改进了 0.241 g/kg、0.231 g/kg 和 0.191 g/kg, MRE 分别平均改进了 7.06%、7.41% 和 5.56%。

(4) 将训练样本根据不同地表类型和是否有云的影响进行分类, BP 神经网络法方案三时(即分晴空和云时)温度和水汽混合比廓线反演结果均最好; CNN 法则方案一时(即不对样本数据进行任何分类)反演结果精度最高。

参考文献:

- [1] Wang Ying, Huang Yong, Huang Siyuan. A preliminary study of the retrieval methods for atmosphere temperature and humidity profiles [J]. *Remote Sensing for Land & Resources*, 2008, 20(1): 23-26. (in Chinese)
- [2] Liu Hui, Dong Chaozhua, Zhang Wenjian. New characteristics of satellite infrared atmospheric detector development over the world [J]. *Meteorological Science and Technology*, 2006, 34(5): 600-605. (in Chinese)
- [3] Qiu Jinhuan, Chen Hongbin, Wang Pucai, et al. A prospect on future atmospheric remote sensing [J]. *Chinese Journal of Atmospheric Sciences*, 2005, 29(1): 131-136. (in Chinese)
- [4] Smith Sr W L, Weisz E, Kireev S V, et al. Dual-regression retrieval algorithm for real-time processing of satellite ultraspectral radiances [J]. *Journal of Applied Meteorology and Climatology*, 2012, 51(8): 1455-1476.
- [5] Guan Yuanhong, Ren Jie, Bao Yansong, et al. Research of the infrared high spectral (IASI) satellite remote sensing atmospheric temperature and humidity profiles based on the one-dimensional variational algorithm [J]. *Trans Atmos Sci*, 2019, 42(4): 602-611. (in Chinese)
- [6] Zhu L, Bao Y, Petropoulos G P, et al. Temperature and humidity profiles retrieval in a plain area from Fengyun-3D/HIRAS sensor using a 1D-VAR assimilation scheme [J]. *Remote Sensing*, 2020, 12(3): 435.
- [7] Chakraborty R, Maitra A. Retrieval of atmospheric properties with radiometric measurements using neural network [J]. *Atmospheric Research*, 2016, 181: 124-132.
- [8] Van Damme M, Whitburn S, Clarisse L, et al. Version 2 of the IASI NH3 neural network retrieval algorithm: near-real-time and reanalysed datasets [J]. *Atmospheric Measurement Technique*, 2017, 10(12): 4905-4914.
- [9] Kolassa J, Gentine P, Prigent C, et al. Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation [J]. *Remote Sensing of Environment*, 2017, 195: 202-217.
- [10] Guan Li, Liu Yang, Zhang Xuehui. Application of artificial neural network algorithm in retrieving atmospheric temperature profiles from hyperspectral infrared data [J]. *Trans Atmos Sci*, 2010, 33(3): 341-346. (in Chinese)
- [11] Liu Yang, Guan Li. Study on the inversion of clear sky atmospheric humidity profiles with artificial neural network [J]. *Meteorological Monthly*, 2011, 37(3): 318-324. (in Chinese)
- [12] Huang P, Guo Q, Han C, et al. An improved method combining ANN and 1D-Var for the retrieval of atmospheric temperature profiles from FY-4A/GIIRS hyperspectral data [J]. *Remote Sensing*, 2021, 13(3): 481.
- [13] Milstein A B, Blackwell W J. Neural network temperature and moisture retrieval algorithm validation for AIRS/AMSU and CrIS/ATMS [J]. *Journal of Geophysical Research: Atmospheres*, 2016, 121(4): 1414-1430.

- [14] Mahngren Hansen D, Laparra V, Nielsen A A, et al. Statistical retrieval of atmospheric profiles with deep convolutional neural networks [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2019, 158: 231-240.
- [15] Seemann S W, Borbas E E, Knuteson R O, et al. Development of a global infrared land surface emissivity database for application to clear sky sounding retrievals from multispectral satellite radiance measurements [J]. *Journal of Applied Meteorology and Climatology*, 2008, 47(1): 108-123.
- [16] Zhu L, Li J, Zhao Y, et al. Retrieval of volcanic ash height from satellite-based infrared measurements [J]. *Journal of Geophysical Research: Atmospheres*, 2017, 122(10): 5364-5379.
- [17] Saunders R, Hocking J, Turner E, et al. An update on the RTTOV fast radiative transfer model (currently at version 12) [J]. *Geoscientific Model Development*, 2018, 11(7): 2717-2737.
- [18] Gambacorta A, Barnett C D. Methodology and information content of the NOAA NESDIS operational channel selection for the cross-track infrared sounder (CrIS) [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(6): 3207-3216.
- [19] Yu P, Shi C, Yang L, et al. A new temperature channel selection method based on singular spectrum analysis for retrieving atmospheric temperature profiles from FY-4A/GIIRS [J]. *Advances in Atmospheric Sciences*, 2020, 37(7): 735-750.
- [20] Wanas N, Auda G, Kamel M S, et al. On the optimal number of hidden nodes in a neural network [C]//Conference Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering, 1998, 2: 918-921.
- [21] Gao Daqi. On structures of supervised linear basis function feedforward three-layered neural networks [J]. *Chinese Journal of Computers*, 1998, 21(1): 80-86. (in Chinese)
- [22] Zhou Hongqiang, Huang Lingling, Wang Yongtian. Deep learning algorithm and its application in optics [J]. *Infrared and Laser Engineering*, 2019, 48(12): 1226004. (in Chinese)
- [23] Xue Shan, Zhang Zhen, Lv Qiongying, et al. Image recognition method of anti UAV system based on convolutional neural network [J]. *Infrared and Laser Engineering*, 2020, 49(7): 20200154. (in Chinese)
- [24] Niu Yaxi, Ji Xiaoping. Image retrieval algorithm based on convolutional neural network [J]. *Computer Engineering and Applications*, 2019, 55(18): 201-206. (in Chinese)
- [25] Zhang J, Chen H, Li Z, et al. Analysis of cloud layer structure in Shouxian, China using RS92 radiosonde aided by 95 GHz cloud radar [J]. *Journal of Geophysical Research*, 2010, 115: D00K30.