

文章编号: 1672-8785(2025)06-0024-10

基于 CARS-CNN-GRU 模型的发动机 尾焰红外光谱浓度求解方法

傅 莉^{1*} 张 昆¹ 孙 旭²

(1. 沈阳航空航天大学自动化学院, 辽宁 沈阳 110136;

2. 中国航发沈阳发动机研究所, 辽宁 沈阳 110015)

摘 要: 针对发动机尾焰组分浓度对红外光谱辐射强度的重要性, 提出了一种高效的红外光谱浓度求解模型, 即结合竞争性自适应重加权采样(Competitive Adaptive Reweighted Sampling, CARS)算法与卷积神经网络(Convolutional Neural Network, CNN)-门控循环单元(Gate Recurrent Unit, GRU)深度学习算法的 CARS-CNN-GRU 模型。该方法通过 CARS 算法筛选关键波长, 获取尾焰成分浓度信息, 然后利用 CNN-GRU 模型对序列数据进行长程依赖分析, 实现多尺度特征提取。仿真结果表明, 与传统模型相比, CARS-CNN-GRU 模型在 H_2O 和 CO_2 浓度求解方面具有更高精度, 其均方根误差(Root Mean Square Error, RMSE)分别降至 0.0014 和 0.0017, R^2 值分别为 0.999 和 0.998, 平均绝对误差(Mean Absolute Error, MAE)分别为 0.0011 和 0.0014。本文提出的 CARS-CNN-GRU 模型在红外光谱浓度求解方面展现出优越的性能, 相较于传统方法具有更高精度、稳定性和可靠性, 为军事和民用航空领域的隐身技术、环境监测以及燃烧效率评估等方面提供了有力支持。

关键词: 发动机尾焰; 红外光谱; CARS-CNN-GRU 模型; 浓度求解

中图分类号: V218 文献标志码: A DOI: 10.3969/j.issn.1672-8785.2025.06.004

A Method for Calculating the Infrared Spectrum Concentration of Engine Tail Flame Based on the CARS-CNN-GRU Model

FU Li^{1*}, ZHANG Kun¹, SUN Xu²

(1. School of Automation, Shenyang Aerospace University, Shenyang 110136, China;

2. Shenyang Engine Research Institute of AECC, Shenyang 110015, China)

Abstract: In view of the importance of the concentration of engine tail flame components to the infrared spectrum radiation intensity, an efficient infrared spectrum concentration solution model is proposed, namely the CARS-CNN-GRU model which combines the competitive adaptive reweighted sampling (CARS) algorithm

收稿日期: 2025-02-27

基金项目: 国家自然科学基金项目(61602321)

作者简介: 傅莉(1968-), 女, 辽宁人, 教授, 博士, 主要从事飞行器隐身测试与智能控制以及磁流变液应用技术研究。

*通讯作者: E-mail: ffulli@163.com

with the convolutional neural network (CNN)-gated recurrent unit (GRU) deep learning algorithm. This method uses the CARS algorithm to select the key wavelengths and obtain the tail flame component concentration information. Then the CNN-GRU model is used to perform long-range dependency analysis on the sequence data to achieve multi-scale feature extraction. Simulation results show that compared with the traditional models, the CARS-CNN-GRU model has higher accuracy in solving H_2O and CO_2 concentrations. Its root mean square error (RMSE) is reduced to 0.0014 and 0.0017, respectively. The R^2 value is 0.999 and 0.998, respectively; the mean absolute error (MAE) is 0.0011 and 0.0014, respectively. The CARS-CNN-GRU model proposed in this paper shows superior performance in solving infrared spectral concentration. Compared with traditional methods, it has higher accuracy, stability and reliability, and provides strong support for stealth technology, environmental monitoring and combustion efficiency evaluation in the military and civil aviation fields.

Key words: engine tail flame; infrared spectrum; CARS-CNN-GRU model; concentration solution

0 引言

发动机尾焰的红外辐射特征是利用红外探测设备对飞行器进行探测、识别、告警、跟踪的重要特征依据^[1]。作为飞机的重要红外辐射源,发动机尾焰的辐射光谱主要取决于其成分与温度。尾焰是飞机发动机喷射的高温、高速气流,其主要成分是 H_2O 、 CO_2 、 N_2 、 CO 等高温气体。其中, H_2O 和 CO_2 是产生红外辐射能量的主要活性成分^[1]。不同的航空煤油燃烧会产生不同浓度的 H_2O 和 CO_2 , 所以对发动机尾焰成分浓度的求解显得十分重要。

随着计算机技术的迅猛进步,通过将机器学习算法与红外光谱技术相结合来实现对混合气体组分浓度的精确检测^[3], 已成为当前研究的重点方向。传统的气体检测手段易受外部条件(如温度、气压变动)的干扰,导致测量结果的稳定性欠佳,通常必须与其他技术配合,方能确保检测目的达成。现阶段,红外光谱被认为是最理想的气体检测方法之一。

因此,红外光谱检测技术引起了该领域学者的极大关注,并在实际领域得到了广泛的应用^[4]。Fang L 等^[5]使用遗传算法在傅里叶变换红外(Fourier Transform Infrared, FTIR)光谱分析中识别未知成分。Evseev V^[6]和 Bharadwaj S P^[7]等利用 HITEMP 数据库和 CDSO 数据库得到高温条件下 CO_2 气体的红外光谱,并讨论了 CO_2 在高温下的光谱特性。余段辉^[8]利用径向基神经网络算法识别了包括 CO 在内

的 5 种气体, Shao L 等^[9]利用 FTIR 技术测量空气并获取了大量的光谱样本。他们建立了经典最小二乘(Classical Least Squares, CLS)和偏最小二乘(Partial Least Squares, PLS)定量模型,以获得光谱中 NH_3 和 CH_4 等气体的浓度信息。Zhang L 等^[10]采用混沌优化进行甲醛和苯浓度估计后,预测误差分别降低了 26.03% 和 16.4%。

卷积神经网络(Convolutional Neural Network, CNN)具有快速的图像特征提取能力,无需数据预处理,因此用于处理光谱数据。Cai Y 等^[11]提出了一种结合 CNN 与长短期记忆(Long Short-Term Memory, LSTM)网络的多气体组分测量方法,但现场环境的干扰会导致低浓度气体结果的预测出现显著误差。

本文基于此,建立了一种 CARS-CNN-GRU 模型。通过 CARS 算法实现对无用变量的去除,并利用 CNN-GRU 模型实现发动机尾焰光谱浓度的精确求解。CNN 对一维光谱数据具有强大的特征提取能力^[12], 而 GRU 网络则弥补了普通循环网络带来梯度消失、梯度爆炸从而难以训练网络的不足,能更好地适应时间序列信号^[13]。二者结合后有望使模型的泛化能力和鲁棒性大大加强。同时,在光谱数据降维方面, CARS 算法对无信息变量进行有效去除的同时,还可以尽可能地减少共线性变量对模型的影响,最终优选出对于求解目标最为关键的变量^[14]。

1 数据获取与建模

1.1 混合气体数据集构建

利用 HITRAN 数据库获取 100 组不同浓度的 H_2O 和 CO_2 混合气体的吸光度光谱并将其作为实验数据。获取的波长区间为 $2\sim 5\ \mu\text{m}$ ，每个光谱包含 3000 个离散数据点(间隔 $0.001\ \mu\text{m}$)。将 80% 数据作为训练集，并将剩下 20% 数据作为测试集。训练集用于训练神经网络，测试集用于评价各个气体浓度求解模型的性能。部分混合气体光谱数据如图 1 所示。

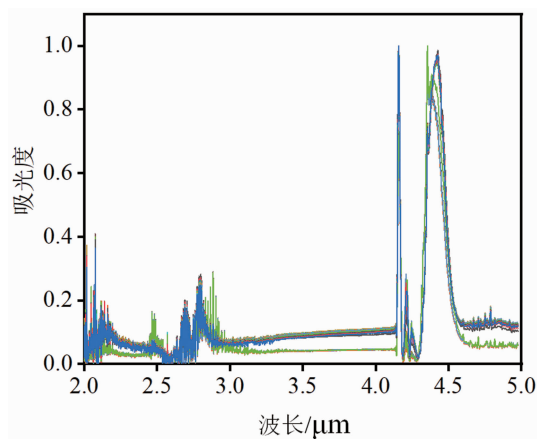


图 1 部分混合气体光谱

1.2 CARS 算法

CARS 算法^[15]是一种用于变量选择的方法，特别适用于光谱数据分析。它能够有效地从大量变量中筛选出最具信息量的特征。CARS 算法的核心在于通过迭代的过程，结合变量的权重调整和逐步剔除不重要变量的策略，来确定最优的变量子集。下面是一个简化版的 CARS 算法流程及其数学描述。

(1) 初始化过程

假设我们有 N 个样本，每个样本包含 p 个变量(特征)。初始时，每个变量的权重 $\omega_i = 1$ ，其中 $i = 1, 2, \dots, p$ 。

(2) PLS 回归

使用 PLS 回归分析，基于当前的权重向量 ω ，计算求解模型。PLS 是一种线性建模技术，特别适用于高维数据。它寻找解释响应变量和求解变量之间关系的潜在线性组合。

(3) 权重更新

计算每个变量的重要性度量，例如变量系数的绝对值 $|\beta_i|$ 。

通常使用式(1)更新变量的权重：

$$\omega_i^{(k+1)} = \omega_i^{(k)} \exp(-\lambda |z_i|) \quad (1)$$

式中， $\omega_i^{(k)}$ 是第 k 次迭代中变量 i 的权重； λ 是调节参数； z_i 是变量 i 的标准化系数。

(4) 竞争性适应性采样

根据更新后的权重，选择一定比例的变量作为“获胜者”。这些变量将保留下来参与下一轮迭代。剩余的“失败者”变量将被淘汰，或者它们的权重将被进一步降低，直到满足停止准则为止。

(5) 迭代过程

上述流程需持续重复执行，直至达成终止条件，比如迭代次数抵达预设上限或者权重变动低于某一临界值。

(6) 最终模型

最后，从保留下来的变量中构建最终的模型。这些变量被认为是数据集中最具代表性和信息量的特征。

CARS 算法通过动态调整样本权重和特征子集，使得重要的特征得到更高的权重，进而提高特征选择和降维的效果。该算法具有较好的鲁棒性和泛化能力。

1.3 CNN

CNN 的设计灵感来源于生物视觉皮层的工作机制，其架构通过交错布置卷积层与池化层，在降低数据规模的同时，有效地捕获和提炼有价值的信息。经全球科研人员持续不懈的探究与改良，CNN 得到了迅猛的更新与发展。该网络包括五个主要部分：输入层、若干卷积层、池化层、全连接层以及输出层。其结构如图 2 所示。

1.4 GRU 网络

GRU 网络属于循环神经网络(Recurrent Neural Network, RNN)家族的一员，其设计理念源自 LSTM 网络的改良。LSTM 网络通过引入具有创新性的输入门控、输出门控以及遗

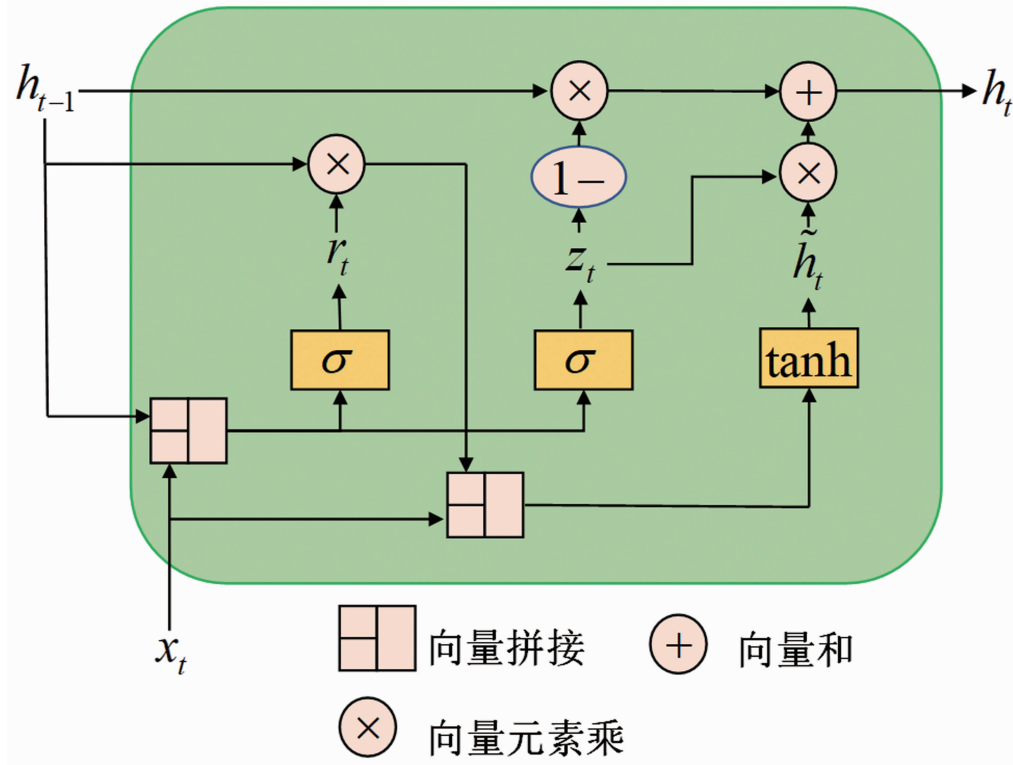


图 4 GRU 节点模型

式中, σ 是逻辑 S 形函数; W_r 是权重向量。 r_t 求解后将用于候选隐藏状态的公式:

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (3)$$

由式(3)可知, r_t 越大, 上一时刻对当前时刻的结合越多。若 r_t 取值为 1, 则表明上一时刻的隐藏状态信息将被完整保留; 若 r_t 取值为 0, 则意味着上一时刻的隐藏状态信息将被完全忽略。因此, 重置门在处理时间序列数据时, 能够有效甄别并捕捉短期内的依赖联系。

更新门用于对记忆进行更新, 其数学公式如下:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (4)$$

更新记忆的表达式如下:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (5)$$

最终的记忆 h_t 忘记传递下来的 h_{t-1} 中的某些信息, 并加入当前节点输入的某些信息。其中, z_t 越接近 1, 代表“记忆”下来的数据越多; z_t 越接近 0, 则代表“遗忘”的数据越多。

1.5 CARS-CNN-GRU 模型建立

在求解混合气体光谱浓度时, 采用 CARS

算法进行特征波段选择, 同时对选择后的数据利用 CNN-GRU 模型作为算法基础, 可以有效提取和利用光谱中的特征信息, 从而提高求解精度。因此, 本节将以上算法与模型结合得到 CARS-CNN-GRU 模型, 并将其与其他算法进行对比。该模型的建模过程如图 5 所示。

CARS-CNN-GRU 模型中包含输入层、CARS 算法、卷积层、池化层、GRU 层、全连接层和输出层。输入混合气体在特征波段的吸光度光谱数据, 利用不同模型对浓度进行求解。

1.6 模型评价指标

引入均方根误差(Root Mean Squared Error, RMSE)、平均绝对误差(Mean Absolute Error, MAE)以及 R^2 决定系数三种误差评价指标, 进而有效地对混合气体浓度求解误差进行评估。RMSE、MAE 以及 R^2 的表达式如下:

$$RMSE(X, h) = \sqrt{\frac{1}{m} \sum_{i=1}^m (h(x_i) - y_i)^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (7)$$

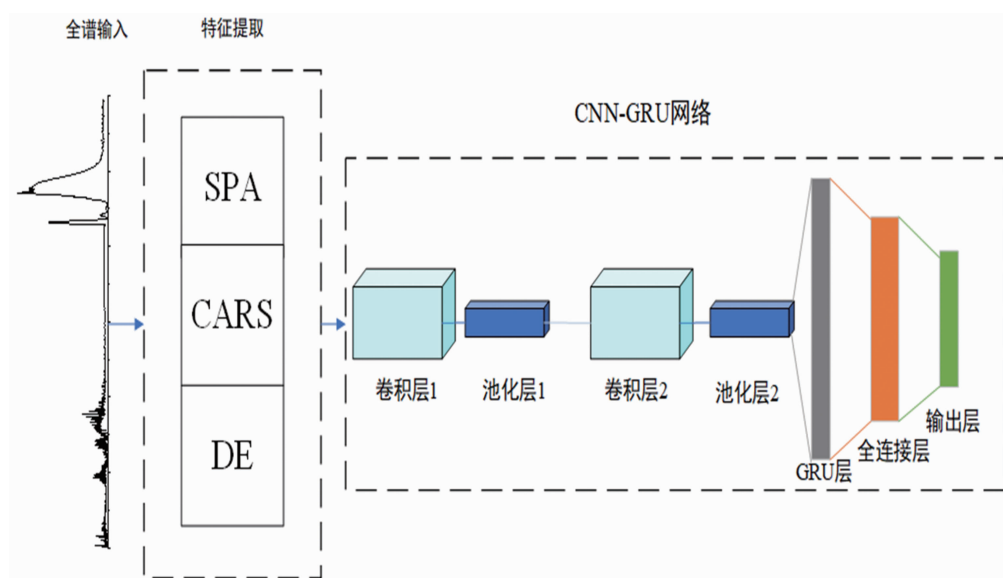


图 5 CARS-CNN-GRU 网络建模图

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (8)$$

式中, y_i 为气体浓度的真实值; $\bar{y}_i$ 为气体浓度的平均值; $\hat{y}_i$ 为气体浓度的求解值; N 为样本数。

2 实验结果分析

2.1 特征波长提取

利用 CARS 算法执行一个 50 轮的变量选择过程(该过程针对混合气体的光谱吸收波段)。随着蒙特卡洛采样迭代次数的增加, 所选波长的数量逐渐减少, 且减速趋于平缓。这一过程展示了波长选择从宽泛到精细的发展趋势。通过比较每次迭代的交叉验证均方根误差(Root Mean Squared Error of Cross-Validation, RMSECV), 整个流程持续进行, 直到找到对应最小 RMSECV 值的迭代, 从而确定了包含最优变量子集的波长。

图 6 为 CARS 特征波段选择的示意图。可以看出, 随着迭代次数的递增, 被选中的波长数目呈现递减趋势, 而 RMSECV 曲线则表现出先波动式下降至最低点随后又逐步回升的态势。这一变化过程表明, 在初期阶段, 通过剔除无关波长变量, RMSECV 值得以降低; 然

而, 在后续阶段, 过度剔除相关变量导致信息缺失, RMSECV 值又随之增加。具体而言, 当 CARS 算法应用于发动机尾焰红外辐射光谱时, 其 RMSECV 值在迭代进行到第 10 次时达到最小值。此时所确定的波长变量数量为 112 个。

为了验证 CARS 算法对模型精度的有效改进, 本文还使用差分进化(Differential Evolution, DE)算法和连续投影算法(Sequential Projection Algorithm, SPA)进行特征波段选取, 并将其与 CARS 算法进行对比验证。

2.2 模型求解结果与分析

将 CARS 特征波段选择后的数据输入训练完成的 CNN-GRU 气体浓度求解模型, 由神经网络求解得到测试集气体样本的浓度。同时, 额外采用 CARS-GRU、SPA-GRU、SPA-CNN-GRU、DE-GRU、DE-CNN-GRU 五种方法对测试集数据进行浓度反演, 然后将结果与 CARS-CNN-GRU 模型的求解结果进行对比。以测试集中所有样本与求解的 RMSE、MAE、 R^2 作为算法性能的评价指标。

利用 CARS-CNN-GRU 模型对 H_2O 和 CO_2 气体浓度的求解结果如图 7(a)和图 7(b)所示。

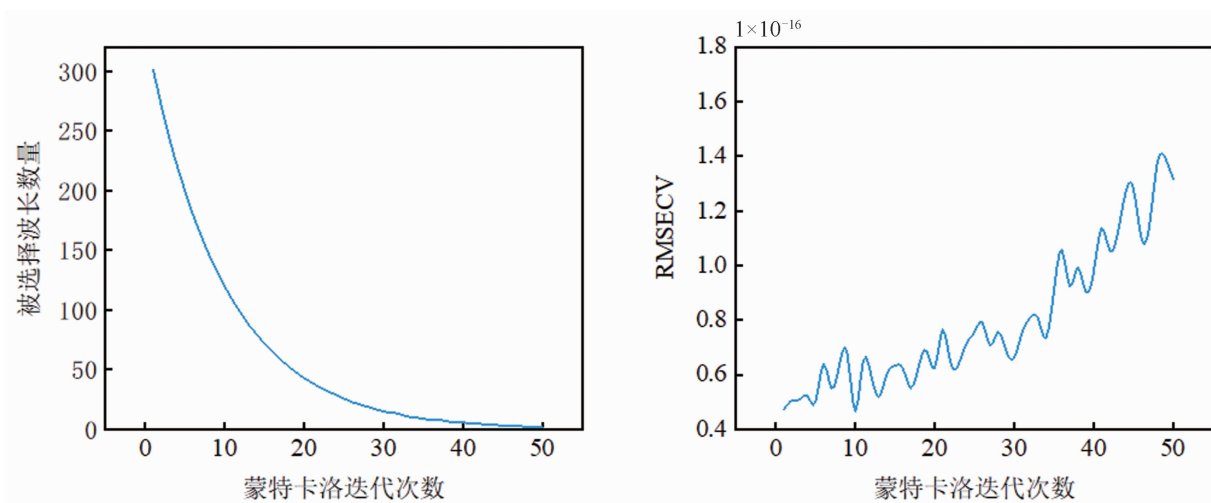
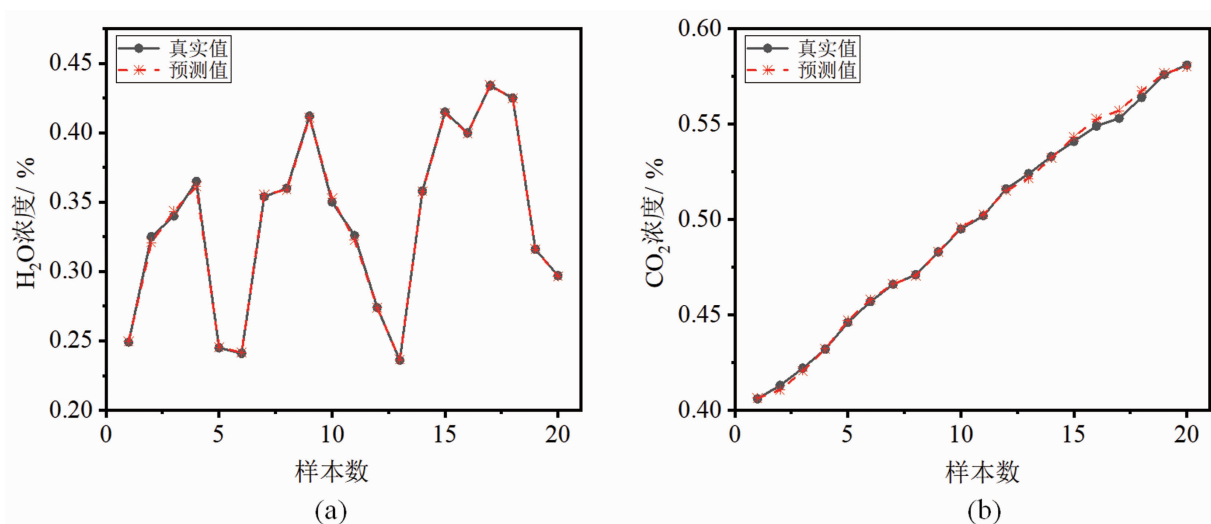


图 6 CARS 特征波段选择的示意图

图 7 混合气体浓度求解结果：(a) H_2O ；(b) CO_2

为了直观体现出本文所提出的 CARS-CNN-GRU 模型的优越性, 现将不同模型的 $RMSE$ 、 MAE 、 R^2 进行对比, 如图 8 所示。

表 1 列出了各算法对 H_2O 和 CO_2 浓度求解的 $RMSE$ 、 MAE 和 R^2 。结合图 8 和表 1 可以得出, 在 $RMSE$ 方面, $\text{CARS-CNN-GRU} < \text{DE-CNN-GRU} < \text{SPA-CNN-GRU} < \text{DE-GRU} < \text{SPA-GRU} < \text{CARS-GRU}$; 在 MAE 方面, $\text{CARS-CNN-GRU} < \text{SPA-CNN-GRU} < \text{DE-CNN-GRU} < \text{DE-GRU} < \text{SPA-GRU} < \text{CARS-GRU}$; 在 R^2 方面, $\text{CARS-CNN-GRU} > \text{SPA-CNN-GRU} > \text{DE-CNN-GRU} > \text{DE-GRU} > \text{SPA-GRU} > \text{CARS-GRU}$ 。这说明 CARS-CNN-GRU

模型的求解误差最小, 求解值与真实值的符合程度最好, 所有样本都得到了较好的浓度求解, 证明 CARS-CNN-GRU 模型具有良好的泛化能力。

CARS-CNN-GRU 模型的气体浓度求解误差低于其他混合模型, 说明利用 CARS 算法对数据特征进行提取后再利用 CNN 算法深度提取特征可以降低误差, 增强 GRU 网络的求解性能。

总的来说, CARS-CNN-GRU 气体浓度求解模型取得的 $RMSE$ 值最小, MAE 值最小, R^2 值最大。因此, CARS-CNN-GRU 是气体浓度求解中求解性能最优的模型。

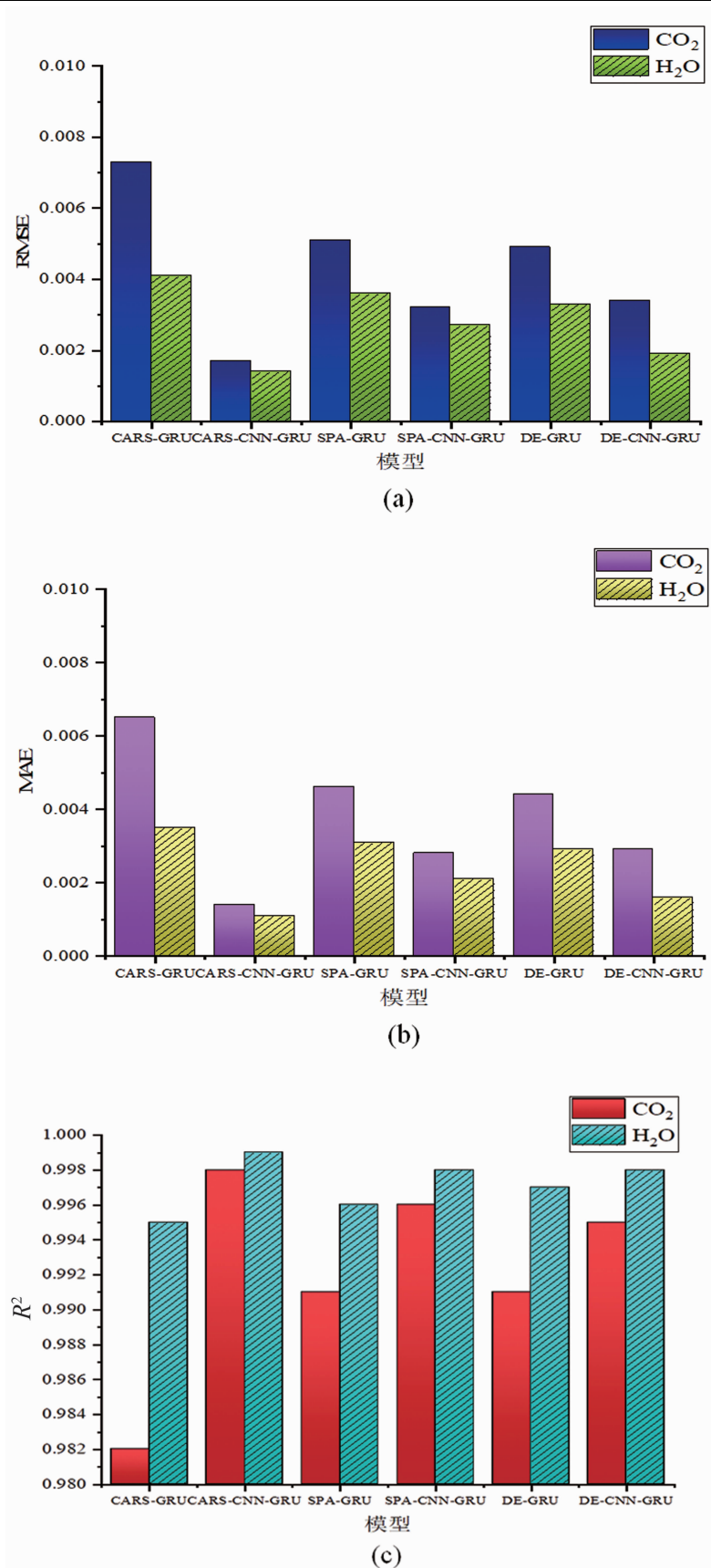
图 8 模型误差指标对比图: (a) RMSE; (b) MAE; (c) R²

表 1 模型求解结果

模型	H ₂ O			CO ₂		
	RMSE	R ²	MAE	RMSE	R ²	MAE
CARS-GRU	0.0041	0.995	0.0035	0.0073	0.982	0.0065
CARS-CNN-GRU	0.0014	0.999	0.0011	0.0017	0.998	0.0014
SPA-GRU	0.0036	0.996	0.0031	0.0051	0.991	0.0046
SPA-CNN-GRU	0.0027	0.998	0.0021	0.0032	0.996	0.0028
DE-GRU	0.0033	0.997	0.0029	0.0049	0.991	0.0044
DE-CNN-GRU	0.0019	0.998	0.0016	0.0034	0.995	0.0029

3 结束语

为了求解出飞机尾焰混合气体成分中的浓度值, 本文在 GRU 网络的基础上, 加入了 CNN 网络, 构建了 CNN-GRU 模型; 同时采用 CARS 算法对数据进行了特征提取, 最终提出了基于 CARS-CNN-GRU 模型的尾焰红外光谱浓度求解方法。同时, 额外采用 CARS-GRU、SPA-GRU、SPA-CNN-GRU、DE-GRU、DE-CNN-GRU 五种方法对测试集数据进行了浓度求解, 并将结果与 CARS-CNN-GRU 模型的预测结果进行了对比。仿真结果表明, CARS-CNN-GRU 模型相较于其他模型在红外光谱浓度求解方面具有更高的精度, 对 H₂O 和 CO₂ 浓度求解的 RMSE 值分别低至 0.0014 和 0.0017, R² 值分别为 0.999 和 0.998, MAE 值分别为 0.0011 和 0.0014。因此, 该模型具有良好的尾焰浓度求解能力。

本文的浓度求解模型目前只适用于高浓度气体的求解, 后续还将进行尾焰中微量气体的求解, 进而扩充该算法的适用条件。

参考文献

- [1] 祝念, 高思莉, 岳娟. 高速飞行目标尾焰红外辐射特性的建模仿真计算 [J]. 红外, 2018, **39**(5): 8-12.
- [2] 于鲲, 丛明煜, 戴文聪. 飞机尾焰的固体粒子红外辐射抑制效果仿真分析 [J]. 光学学报, 2020, **40**(21): 191-204.
- [3] Khoury Y E, Gebelin M, Marcou G, et al. Rapid discrimination of neuromyelitis optica spectrum disorder and multiple sclerosis using machine learning on infrared spectra of sera [J]. *International Journal of Molecular Sciences*, 2022, **23**(5): 2791.
- [4] Liang H, Long Y, Liu G. Qualitative and quantitative studies of multicomponent gas by CNN-KP-CA-RF model [J]. *Vibrational Spectroscopy*, 2024, **130**: 103647.
- [5] Fang L, Junde W. Using genetic algorithm to identify completely unknown system in FTIR spectra analysis [J]. *Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering*, 2004, **39**(6): 1525-1533.
- [6] Evseev V, Fateev A, Clausen S. High-resolution transmission measurements of CO₂ at high temperatures for industrial applications [J]. *Journal of Quantitative Spectroscopy & Radiative Transfer*, 2012, **113**(17): 2222-2233.
- [7] Bharadwaj S P, Modest M F. Medium resolution transmission measurements of CO₂ at high temperature [J]. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 2007, **73**(2): 329-338.
- [8] 余段辉. 基于红外光谱的气体识别与浓度检测算法研究 [D]. 成都: 电子科技大学, 2018.
- [9] Shao L, Griffiths P R, Chu P M, et al. Quantitative vapor-phase infrared spectrometry of ammonia [J]. *Applied Spectroscopy*, 2006, **60**(3): 254-260.
- [10] Zhang L, Tian F, Liu S, et al. Chaos based neural network optimization for concentration estimation of indoor air contaminants by an electronic nose [J]. *Sensors and Actuators A: Physical*, 2013, **189**: 161-167.

- [11] Cai Y, Xu G, Yang D, et al. On-line multi-gas component measurement in the mud logging process based on Raman spectroscopy combined with a CNN-LSTM-AM hybrid model [J]. *Analytica Chimica Acta*, 2023, **1259**: 341200.
- [12] 刘宗溢, 张彩虹, 蒋健康, 等. 基于拉曼光谱结合 CNN-LSTM 深度学习方法的铁皮石斛总黄酮含量快速检测研究 [J]. *光谱学与光谱分析*, 2024, **44**(4): 1018-1024.
- [13] 彭朝琴, 李奇聪, 陈娟, 等. 基于 GRU 和改进注意力机制的多信息融合的 EMA 故障诊断方法 [J/OL]. <https://bhxb.buaa.edu.cn/bhbk/cn/article/doi/10.13700/j.bh.1001-5965>. 2023. 0584, 2023.
- [14] 李江波, 彭彦昆, 陈立平, 等. 近红外高光谱图像结合 CARS 算法对鸭梨 SSC 含量定量测定 [J]. *光谱学与光谱分析*, 2014, **34**(5): 1264-1269.
- [15] Li H D, Liang Y Z, Xu Q S, et al. Key wavelengths screening using competitive adaptive re-weighted sampling method for multivariate calibration [J]. *Analytica Chimica Acta*, 2009, **648**: 77-84.
- [16] 陈恩帅, 茅大钧, 陈思勤, 等. 基于双向 LSTM-Attention 模型的火电厂负荷预测研究 [J]. *电力科技与环保*, 2024, **40**(4): 380-387.