

DOI:10.19651/j.cnki.emt.2518398

# 基于火焰成像的 NO<sub>x</sub> 排放混合预测模型 SADAЕ-MSViL<sup>\*</sup>

常胜君<sup>1,2</sup> 郝润芳<sup>1,2</sup> 程永强<sup>1,3</sup> 杨琨<sup>1,2</sup> 白云鹏<sup>1,3</sup>(1. 人工智能微纳传感山西省重点实验室 太原 030600; 2. 太原理工大学集成电路学院 太原 030600;  
3. 太原理工大学电子信息工程学院 太原 030600)

**摘要:** 火电厂生物质燃烧过程中产生的 NO<sub>x</sub> 对环境造成严重污染,准确预测 NO<sub>x</sub> 排放对于降低环境污染至关重要。基于传统数据驱动方法建立的 NO<sub>x</sub> 排放预测模型深层特征信息提取不充分,鲁棒性差。针对现存问题提出了基于火焰成像的 NO<sub>x</sub> 排放混合预测模型 SADAЕ-MSViL。首先,在对抗降噪自动编码器中引入自注意力机制,实现图像深层特征提取,有效去除噪声干扰。其次,设计尺度为 8 和 16 组合的多尺度特征融合机制,充分捕获不同尺度下图像块的火焰频域信息。最后,通过改进 Linformer 并融合门控低秩注意力机制,在保证模型运行效率的同时提升 NO<sub>x</sub> 排放预测精度。实验结果表明,该模型  $R^2$  达到 0.98, RMSE 为 3.0, 预测精度优于其他模型,展现出较高的鲁棒性和可靠性。

**关键词:** NO<sub>x</sub> 排放预测; 火焰温度图像; 自注意力; 多尺度神经网络

**中图分类号:** TN911.73; TK223 **文献标识码:** A **国家标准学科分类代码:** 470.10

## SADAЕ-MSViL: A hybrid prediction model for NO<sub>x</sub> emissions based on flame imaging

Chang Shengjun<sup>1,2</sup> Hao Runfang<sup>1,2</sup> Cheng Yongqiang<sup>1,3</sup> Yang Kun<sup>1,2</sup> Bai Yunpeng<sup>1,3</sup>

(1. Shanxi Provincial Key Laboratory of Artificial Intelligence Micro-Nano Sensing, Taiyuan 030600, China;

2. School of Integrated Circuits, Taiyuan University of Technology, Taiyuan 030600, China;

3. School of Electronic Information Engineering, Taiyuan University of Technology, Taiyuan 030600, China)

**Abstract:** NO<sub>x</sub> produced during biomass combustion in thermal power plants causes serious environmental pollution. Accurate prediction of NO<sub>x</sub> emissions is crucial to reducing environmental pollution. The NO<sub>x</sub> emission prediction model established based on traditional data-driven methods does not extract deep feature information sufficiently and has poor robustness. To address the existing problems, a hybrid prediction model for NO<sub>x</sub> emissions based on flame imaging, SADAЕ-MSViL, is proposed. First, a self-attention mechanism is introduced into the adversarial denoising autoencoder to extract deep features of the image and effectively remove noise interference. Secondly, a multi-scale feature fusion mechanism with a combination of scales of 8 and 16 is designed to fully capture the flame frequency domain information of image blocks at different scales. Finally, by improving Linformer and integrating the gated low-rank attention mechanism, the NO<sub>x</sub> emission prediction accuracy is improved while ensuring the operating efficiency of the model. Experimental results show that the  $R^2$  of the model reaches 0.98 and the RMSE is 3.0. The prediction accuracy is better than other models, showing high robustness and reliability.

**Keywords:** NO<sub>x</sub> emission prediction; temperature images; self-attention; multiscale visual neural networks

## 0 引言

燃烧过程中产生的氮氧化物(nitrogen oxides, NO<sub>x</sub>) 对环境和人类健康造成严重危害。因此,减少 NO<sub>x</sub> 排放成

为全球环保的重要目标之一<sup>[1]</sup>。随着碳排放对气候变化的影响日益加剧,碳中和已成为全球减排努力的核心方向。准确预测 NO<sub>x</sub> 排放量,不仅有助于减少污染物的释放,还能在碳排放控制和碳中和目标实现过程中,推动绿色低碳

收稿日期:2025-03-20

\* 基金项目:山西省重点研发计划项目(202102030201012)、山西省省筹资金资助回国留学人员科研项目(2024-047)资助

经济的发展。

火电厂 NO<sub>x</sub> 排放预测方法主要分为仪器测量、机理模型和数据驱动模型。仪器测量操作复杂且易受环境影响,难以准确测量<sup>[2]</sup>;机理模型基于燃烧公式和经验参数,但因燃烧过程复杂,实际应用困难<sup>[3]</sup>。随着人工智能发展,数据驱动方法广泛应用。研究人员利用深度学习模型建立锅炉热力参数与其燃烧产生 NO<sub>x</sub> 排放量之间的关系<sup>[4-6]</sup>。例如,Liu 等<sup>[7]</sup>利用互信息(mutual information, MI)筛选出 NO<sub>x</sub> 排放关键影响因素并利用长短时记忆模型预测 NO<sub>x</sub> 排放量。

然而,火焰成像相比热力参数,能够捕捉火焰微观特性,提升模型预测精度。周裕等<sup>[8]</sup>利用双尺度去噪的 CNN 模型识别火焰图像,Li 等<sup>[9]</sup>通过火焰自由基特征值识别生物质燃料。传统 CNN 依赖人工特征(如灰度直方图、纹理特征)和大量标记数据,难以全面捕捉火焰图像信息,限制了泛化能力。自动编码器(AE)<sup>[10]</sup>无需大量标记数据,可自动提取特征,减少人工依赖。Tang 等<sup>[11]</sup>利用 AE 建立火焰图像与 NO<sub>x</sub> 排放的非线性映射。但传统 AE 处理高维图像效率低且易受噪声干扰,需增强特征提取和抗噪性能。此外,传统模型预测 NO<sub>x</sub> 排放准确率较低。王宝华等<sup>[12]</sup>使用卷积自编码(convolutional autoencoder, CAE)提取模型预测 NO<sub>x</sub>;Qin 等<sup>[13]</sup>通过在线增量学习更新迭代提升模型精度。但传统模型缺乏多尺度信息融合能力,局部与全局信息整合不足,限制了预测性能。因此,亟需引入多尺度特征融合机制,提升模型精度和泛化能力。

针对上述问题,本文提出一种基于火焰成像的 NO<sub>x</sub> 排放混合预测模型,该模型在对抗降噪自动编码器(adversarial denoising autoencoders, ADAE)<sup>[14]</sup>中引入自注意力机制(self-attention adversarial denoising autoencoders, SADAe)进行特征提取,并设计多尺度视觉神经网络(multi scale Vision Linformer, MSViL)实现 NO<sub>x</sub> 排放预测。实验使用的数据来自英国工业规模的测试设施,并利用不同燃烧条件下的火焰图像评估了 NO<sub>x</sub> 排放预测方法。本文贡献如下:

- 1) 提出 SADAe 模型,通过自注意力机制去除火焰图像噪声并从高维火焰图像中充分提取深度特征,解决模型鲁棒性差的问题。
- 2) 设计 MSViL 模型融合不同尺度局部和全局火焰频域信息,确定模块最优尺度组合与结构参数。
- 3) 改进 Linformer 并融合门控低秩注意力机制,降低计算复杂度,同时利用门控权重动态调节投影矩阵,实现了精度与效率的平衡。

## 1 理论方法

本文提出的 SADAe-MSViL 模型由两部分组成:SADAe 模型和 MSViL 模型。在模型训练之前,首先对火焰图像进行预处理;随后,SADAe 模型提取火焰图像的深

层特征;最后,MSViL 模型利用这些特征,实现高精度的 NO<sub>x</sub> 浓度预测,如图 1 所示。

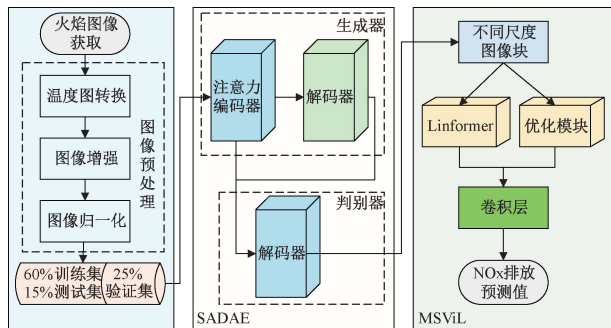


图 1 NO<sub>x</sub> 排放预测流程

Fig. 1 NO<sub>x</sub> emission prediction process

### 1.1 自注意力对抗去噪自动编码器

SADAe 基于 ADAE 模型在编码器中加入了 SA 机制用于提取图像的深度特征,如图 2 所示。

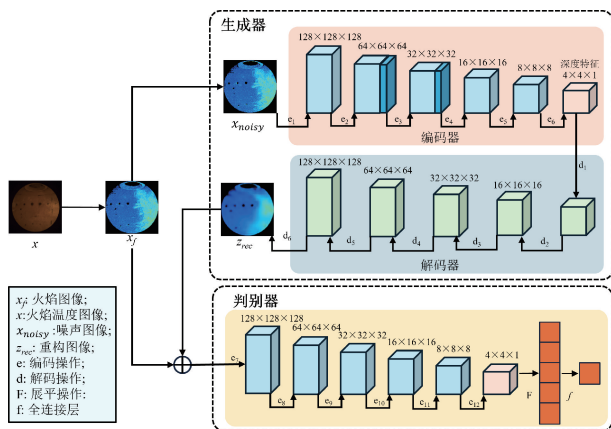


图 2 SADAe 结构图

Fig. 2 SADAe structure diagram

#### 1) 理论基础

ADAE 模型中为了避免了输入信息的简单复制,加入噪声可以表示为:

$$x_{noisy(i)} = x + \delta\tau \quad (1)$$

其中, $\delta$  表示噪声水平; $\tau$  表示服从正态分布的随机变量。 $x_{noisy}$  作为编码器输入,通过解码器输出为重构图像  $z_{rec}$ ,得到最优重构图像误差  $L_{img}$ :

$$L_{img} = \frac{1}{n} \sum_{i=1}^n \| z_{rec(i)} - x_i \| \quad (2)$$

针对陷入局部最优问题,ADAE 中生成对抗网络(generative adversarial networks, GAN)<sup>[15]</sup>有效解决了压缩过程中因过度关注全局特征而忽略局部细节的问题,避免了关键信息丢失,从而显著提升了图像重构质量。GAN 机制如式(3),结构如图 3 所示。

$$L_{GAN} = \min_G \max_D \{ E_{P(G(\alpha))} [\log[1 - D(G(\alpha))]] + E_{P(\beta)} [\log D(\beta)] \} \quad (3)$$

其中,  $E(\cdot)$  表示期望;  $D(\cdot)$  表示判别器输出;  $G(\cdot)$  表示生成器输出;  $P(G(\alpha))$  和  $P(\beta)$  分别表示噪声数据的概率分布和真实数据的概率分布。

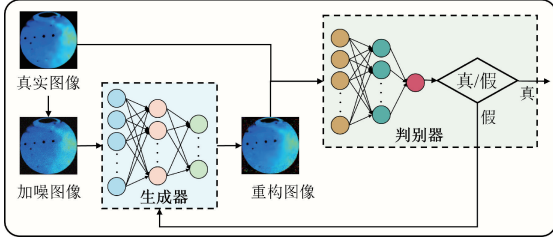


图 3 GAN 结构图

Fig. 3 GAN structure diagram

## 2) 改进模型

温度图像重构中, ADAЕ 存在高频细节损失和难以捕捉不同区域间的相关性等问题。SA 的引入可以有效弥补

这些缺陷, 通过增强全局信息建模、增加生成多样性以及优化训练稳定性。本文在编码器  $e_1$  和  $e_2$  后加入 SA 模块, 在降维后的特征图上计算注意力权重, 强制关注火焰图像中具有高相关性的跨区域信息。在  $e_1$  后 SA 模块聚焦相邻像素的温度梯度变化, 保留高频火焰细节, 在  $e_2$  后 SA 模块捕捉火焰核心与边缘区域的动态交互, 提升全局建模能力。这一改进不仅弥补了 GAN 在全局信息建模上的不足, 还显著增强了模型对火焰动态特征的捕捉能力。4.2 节中将讨论 SA 在 SADAЕ 中的位置。在模型架构中,  $x_{noisy}$  经过一系列编码操作后得到维度为  $4 \times 4 \times 1$  的深层特征  $x_{dpf}$ , 随后采用与编码器相反操作, 经过全部解码操作后得到重构图像  $z_{rec}$ , 用于判别器与真实图像进行比较并判别真假, 全部参数由表 1 可见。通过充分捕获初次和二次降维后的微小特征, SA 模块显著提升了模型的重构精度、鲁棒性和泛化能力。

表 1 模型参数配置

Table 1 Model parameter configuration

模型	上采样层	卷积层	激活函数	池化层	输出
输入图像	—	—	—	—	(256, 256, 3)
编码器 $e_1$	—	$q=128; k=3; s=1$	ReLU	$k=2; s=2$	(128, 128, 128)
编码器 $e_2$	—	$q=64; k=3; s=1$	ReLU	$k=2; s=2$	(64, 64, 64)
编码器 $e_3$	—	$q=32; k=3; s=1$	ReLU	$k=2; s=2$	(32, 32, 32)
编码器 $e_4$	—	$q=16; k=3; s=1$	ReLU	$k=2; s=2$	(16, 16, 16)
编码器 $e_5$	—	$q=8; k=3; s=1$	ReLU	$k=2; s=2$	(8, 8, 8)
编码器 $e_6$	—	$q=1; k=3; s=1$	ReLU	$k=2; s=2$	(4, 4, 1)
解码器 $d_1$	$k=2$	$q=8; k=3; s=1$	ReLU	—	(8, 8, 8)
解码器 $d_2$	$k=2$	$q=16; k=3; s=1$	ReLU	—	(16, 16, 16)
解码器 $d_3$	$k=2$	$q=32; k=3; s=1$	ReLU	—	(32, 32, 32)
解码器 $d_4$	$k=2$	$q=64; k=3; s=1$	ReLU	—	(64, 64, 64)
解码器 $d_5$	$k=2$	$q=128; k=3; s=1$	ReLU	—	(128, 128, 128)
解码器 $d_6$	$k=2$	$q=3; k=3; s=1$	Sigmoid	—	(256, 256, 3)
判别器 $e_7$	—	$q=128; k=4; s=2$	LeakyReLU	—	(128, 128, 128)
判别器 $e_8$	—	$q=64; k=4; s=2$	LeakyReLU	—	(64, 64, 64)
判别器 $e_9$	—	$q=32; k=4; s=2$	LeakyReLU	—	(32, 32, 32)
判别器 $e_{10}$	—	$q=16; k=4; s=2$	LeakyReLU	—	(16, 16, 16)
判别器 $e_{11}$	—	$q=8; k=4; s=2$	LeakyReLU	—	(8, 8, 8)
判别器 $e_{12}$	—	$q=1; k=4; s=2$	LeakyReLU	—	(4, 4, 1)
展平	—	—	Sigmoid	—	(1)

注:  $q$  为滤波器数量;  $k$  为卷积核尺寸;  $s$  为 stride

为了进一步提升 SADAЕ 生成图像质量, 引入多尺度结构相似性损失 (multi scale-structure similarity index measure, MS-SSIM), 使模型不仅能在像素级别上逼近原始图像, 还能有效捕捉图像的多尺度结构特征, 确保生成图像在视觉上更加自然和谐。如式 (4) 所示。

$$L_{MS-SSIM} = 1 - \prod_{j=1}^M [\text{SSIM}(x^j, y^j)]^{\gamma_j} \quad (4)$$

其中,  $x^j, y^j$  分别是输入图像和生成图像在第  $j$  级别的下采样版本,  $\gamma_j$  是不同尺度的加权参数。总体损失  $Loss$  公式为:

$$Loss = \lambda_1 L_{img} + \lambda_2 L_{GAN} + \lambda_3 (L_{MS-SSIM}) \quad (5)$$

## 1.2 多尺度视觉神经网络

本文针对 ViT (vision transformer)<sup>[16]</sup> 中的多头自注意力 (multi-head self-attention, MHSA) 机制在高分辨率火

焰图像分析中计算成本高、内存占用大的问题,引入多尺度低秩 Linformer 注意力机制,设计了 MSViL 模型。该模型通过优化注意力计算结构,显著降低了计算资源消耗。

ViL(vision linformer)模型的结构如图 4 所示。经过 SADA 模型重构的图像  $z_{rec}$  被分割为  $N$  个  $P \times P$  的非重叠图像块:

$$N = \frac{H \times W}{P \times P} \quad (6)$$

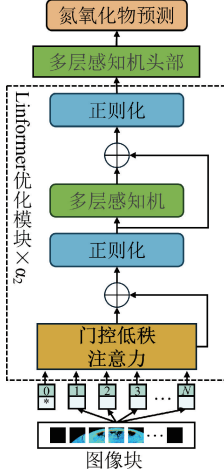


图 4 ViL 模型结构图

Fig. 4 ViL model structure diagram

每个图像块展平为一维向量,经过线性变换映射到维度  $D$ ,得到嵌入块  $z_i \in \mathbb{R}^D$ 。此外,整体图像  $X_{ES}$  表示为:

$$X_{ES} = [z_{cls}, z_1, z_2, \dots, z_n] + E_{pos} \quad (7)$$

其中,  $E_{pos} \in \mathbb{R}^{(N+1) \times D}$  为位置编码,添加位置信息。通过可学习的线性投影矩阵对键(key,  $K$ )和值(value,  $V$ )进行降维,使每个查询(query,  $Q$ )能够在低维空间中更高效地提取关键信息。本文对键和值进行非线性降维,同时引入门控机制以适应不同输入:

$$K' = \sigma_{sig}(K \cdot f_{\theta}(X_{ES})) \odot K \quad (8)$$

$$V' = \sigma_{sig}(V \cdot g_{\theta}(X_{ES})) \odot V \quad (9)$$

其中,  $K'$  和  $V'$  为非线性降维后的键和值;  $f_{\theta}(\cdot)$  和  $g_{\theta}(\cdot)$  是自适应的非线性映射函数;  $\sigma_{sig}(\cdot)$  代表 sigmoid 激活函数用于生成门控权重;  $\odot$  表示元素级乘法。这种改进可以使投影矩阵随输入动态变化,同时门控机制能够抑制不必要的信息。

除此之外,本文将传统注意力机制改为自适应归一化版本,并加入温度因子  $\tau$  以及门控参数  $G$  来增强灵活性,低秩注意力计算公式为:

$$AT(Q, K', V') = G \odot \text{Norm}\left(\frac{QK'^T}{\sqrt{d_k} \cdot \tau(X_{ES})}\right) V' \quad (10)$$

其中,对于第  $l$  个 Linformer 层,输入为  $Z_{rec}^{(l)} \in \mathbb{R}^{B \times (n_l+1) \times D}$ ,输出更新为  $Z_{rec}^{(l+1)}$ ,公式为:

$$head_i = \text{Attention}(Q_i, V_i, V'_i) \quad (11)$$

$$\text{MHSA}(X) = [\text{head}_1, \text{head}_2, \dots, \text{head}_{\alpha_2}] W_O \quad (12)$$

$$Z_{rec}^{(l+1)} = \text{LN}(Z_{rec}^{(l)} + \text{MHSA}(Z_{rec}^{(l)})) \quad (13)$$

式中:  $\alpha_2$  为注意力头的数量;  $W_O \in \mathbb{R}^{(h \cdot d_k) \times D}$  为投影矩阵。通过残差连接和多层感知机(multilayer perceptron, MLP)输出  $Z_{MLP}^{(l+1)} \in \mathbb{R}^{B \times n_l \times D}$ ,公式为:

$$Z_{MLP}^{(l+1)} = \text{LN}(Z_{rec}^{(l+1)} + \text{MLP}(Z_{rec}^{(l+1)})) \quad (14)$$

$$\text{MLP}(Z_{rec}^{(l+1)}) = f(\sigma(f(Z_{rec}^{(l+1)}))) \quad (15)$$

其中,  $f(\cdot)$  代表全连接操作;  $\sigma(\cdot)$  代表 ReLU 激活函数。

本文引入了多尺度融合机制,突破了传统单尺度特征提取方法的局限性。将不同尺寸的图像块输入到 MSViL 模型中,并对输出结果进行拼接融合,如图 5 所示。结合低秩注意力机制和多尺度融合机制,模型提取局部和全局不同频域特征,实现了预测精度与计算效率的平衡。

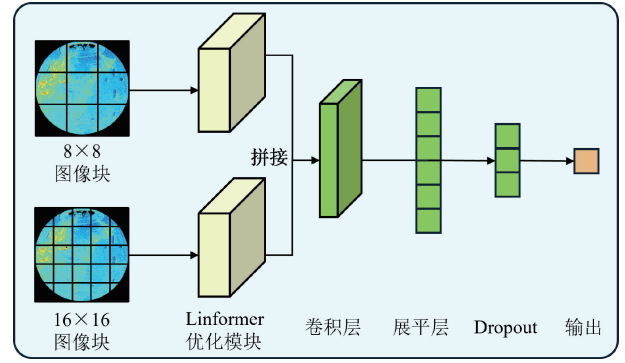


图 5 多尺度融合

Fig. 5 Multi-scale fusion

对每个分支的输出  $Z_{MLP(i,j)}^{(l+1)}$ , ( $i = 1, 2; j = 1, 2, \dots, n_i$ ),  $n_i \in \mathbb{R}^{N \times 1}$  为第  $i$  个尺度下的图像块数量,在序列维度上取平均得到  $\bar{Z}_i$ :

$$\bar{Z}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Z_{MLP(i,j)}^{(l+1)}, \bar{Z}_i \in \mathbb{R}^{B \times d} \quad (16)$$

然后将所有尺度的特征拼接通过 MLP 输出结果  $Y \in \mathbb{R}^{B \times 1}$ :

$$Z_{con} = \text{Concat}(\bar{Z}_1, \bar{Z}_2), Z_{con} \in \mathbb{R}^{B \times (2 \cdot d)} \quad (17)$$

$$Y = \sigma(Z_{con} W + b) \quad (18)$$

## 2 实验研究

### 2.1 实验数据

在英国碳捕获和储存研究中心的 250 kWth 燃烧测试设施中,进行了氧生物质燃烧实验。图 6 展示了 CTF 及其火焰成像系统的布局示意图。火焰成像系统由光学探头、摄像机和计算机组成。光学探头是一个带水冷夹套的刚性内窥镜,具有 90°视野。摄像机采用工业 RGB CMOS 传感器,分辨率为 1 216 pixel × 1 936 pixel,最高帧率可达 900 fps。计算机运行专用软件处理图像,整个系统安装在



炉顶的观察口上。

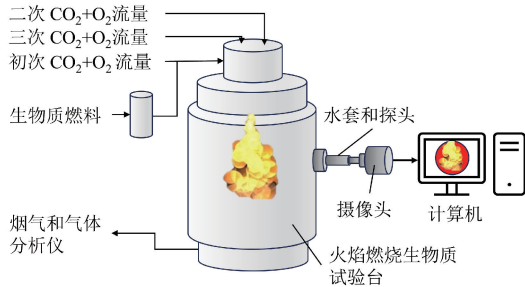


图 6 火焰成像系统

Fig. 6 Flame imaging system

实验选用粉碎后的白木作为生物质燃料,并以 CO<sub>2</sub> 和 O<sub>2</sub> 的混合物模拟富氧燃烧环境。通过调节二次和三次气流(CO<sub>2</sub> + O<sub>2</sub>)、氧气浓度、一次气流与总气流比例以及二次与三次气流比例,创建了 10 种不同的操作条件。在每种稳定状态下,使用火焰成像系统采集了约 150~170 张图像,总数据收集持续约 4 min,共获得 1 660 张图像。将每种条件下的图像数据按 1:4 的比例划分,其中 1/4 作为验证集并取出,剩余 3/4 按 4:1 的比例划分为训练集和测试集。在该比例下的验证集图片可以对门控机制、低秩维度和不同尺度等改进模块进行充分对比和评估。

为增加样本多样性,对输入数据集进行了翻转和随机裁剪等数据增强操作。气体分析仪用于测量烟气中的 NO<sub>x</sub> 含量,为模型开发提供了关键排放基准。表 2 汇总了所有测试条件及相关操作参数,图 7 展示了不同条件下的火焰图像和通过双色方法<sup>[17]</sup>计算的温度图像。

表 2 测试条件总结

Table 2 Summary of test conditions

测试 条件	O <sub>2</sub> / (Vol%)	1st/T/ (%)	2nd/ 3rd	Mean/ (mg·m <sup>-3</sup> )	STD/ (mg·m <sup>-3</sup> )
1	21	24	4	339.45	4.65
2	23	24	4	385.95	9.3
3	25	24	4	367.35	9.3
4	27	24	4	362.70	9.3
5	27	21	4	311.52	9.3
6	27	27	4	376.65	9.3
7	27	30	4	437.10	9.3
8	27	24	1	390.60	18.6
9	27	24	2	339.45	13.95
10	27	24	3	348.75	9.3

1st/T 为一次流量与总流量比;2nd/3rd 为二次流量与三次流量的分流比;Mean 为烟气中 NO<sub>x</sub> 平均浓度;STD 为 NO<sub>x</sub> 浓度标准差。

2.2 评价指标

为了全面评估模型的预测性能,本文采用了 3 个评估指

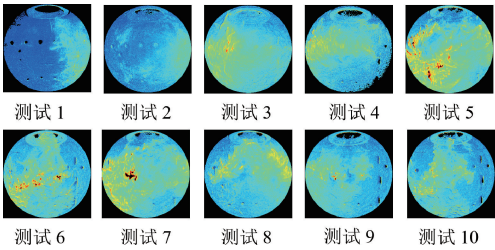


图 7 不同测试条件下的火焰和温度图像

Fig. 7 Flame and temperature images under different test conditions

标。平均绝对误差(mean absolute deviation, MAE)、均方根误差(root mean square error, RMSE)和决定系数(coefficient of determination, R<sup>2</sup>)用于评估模型性能,定义为:

MAE = 1/n \* sum\_{i=1}^n |Y\_{pre(i)} - Y\_{true(i)}|, (19)

RMSE = sqrt(1/n \* sum\_{i=1}^n (Y\_{pre(i)} - Y\_{true(i)})^2), (20)

R^2 = 1 - (sum\_{i=1}^n (Y\_{pre(i)} - Y\_{true(i)})^2) / (sum\_{i=1}^n (Y\_{true(i)} - Y\_bar\_{true(i)})^2), (21)

其中, Y\_{true(i)} 表示真实值; Y\_{pre(i)} 表示预测值; Y\_bar\_{true} 是数据集中所有真实值的均值。

3 模型训练

本文提出的模型在 Windows 11 系统上使用 Python 编写。在配备 Intel i9-13980HX 处理器和 NVIDIA GeForce RTX 4090 的计算机上进行训练和测试。

模型由 SADA-E 和多尺度 ViL 两部分组成。图 8 展示了 SADA-E 训练过程中损失函数的变化趋势。前 30 次迭代中,训练损失和验证损失均快速下降,表明模型能够有效捕捉数据的主要特征;30 至 100 次迭代阶段,损失下降速度显著减缓并逐渐趋于稳定;当迭代次数超过 100 次后,随着训练继续,损失逐渐增加,表明模型开始过拟合。为避免过拟合并降低计算成本,最终将 SADA-E 的训练迭代次数设为 100,以便后续多尺度 ViL 模型的训练。

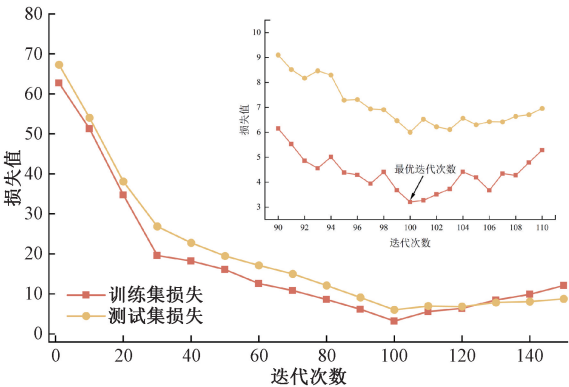


图 8 不同输入条件下的损失比较

Fig. 8 Comparison of losses under different input conditions

将 SADAЕ 生成的火焰温度图像输入多尺度 ViL 模型中进行 NO<sub>x</sub> 排放预测。图 9 显示,预测值与实际测量值的分布曲线高度一致, $R^2=0.98$ ,表明模型对数据的拟合能力较强,能够更有效地捕获 NO<sub>x</sub> 排放过程的关键特征。然而,尽管整体拟合效果良好,仍存在一定误差,主要来源于输入数据的测量误差和实验环境误差。

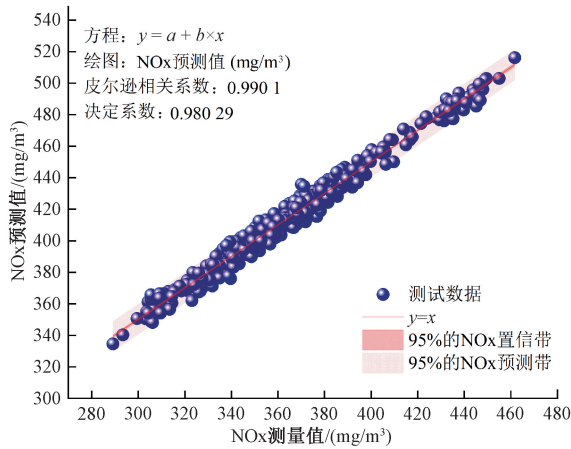


图 9 NO<sub>x</sub> 排放量的测量值与预测值对比  
Fig. 9 Comparison between measured and predicted values of NO<sub>x</sub> emissions

4 结果与讨论

4.1 SADAЕ 模型噪声水平

为了避免模型对图像进行简单的复制而非重构,同时提升模型在噪声干扰下的鲁棒性,在 SADAЕ 训练中引入不同水平的高斯噪声  $\delta$ 。本文通过调整噪声水平进行了实验,结果如图 10 所示。随着噪声水平的增加,SADAЕ 模型的预测精度呈现先下降后上升的趋势。当噪声水平  $\delta=0.1$  时,模型预测精度达到最低点,训练精度 MAE=1.5,测试精度 MAE=2.22。由此可知,适当的噪声水平能够有效提升模型性能,最终确定最佳噪声水平为  $\delta=0.1$ 。

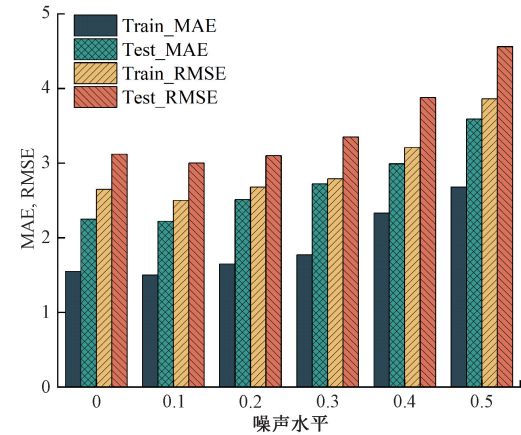


图 10 SADAЕ 在不同噪声水平下的表现  
Fig. 10 Performance of SADAЕ at different noise levels

4.2 SADAЕ 中 SA 的位置优化与影响

SA 所在位置对于特征信息提取有不同的作用,合理地选择 SA 的数量和位置,不仅能够有效提升模型对全局与局部特征的捕捉能力,还能避免计算资源浪费和模型运算负担过重的问题。由表 3 可知未添加 SA 时,虽然模型训练较快但是精度很低,凸显 SA 对于 SADAЕ 性能提升有较大影响。SA 在 SADE 中最佳位置应加在编码器  $e_1$  和  $e_2$  后面,此时 SADAЕ 模型的 MSE 和 RMSE 达到最低,且训练时长相差不多。上述仅在  $e_1, e_2$  和  $e_3$  后加入 SA 是因为其他编码器通道数少于 32,难以充分发挥 SA 的作用,还可能引入冗余的计算开销。

表 3 SA 在 SADE 模型中位置比较  
Table 3 Comparison of the position of SA in the SADE model

SA 位置	MAE	RMSE	Time/s
未添加 SA	1.25	1.32	798
$e_1$ 后	0.94	0.94	853
$e_2$ 后	1.03	0.97	847
$e_3$ 后	1.15	1.02	833
$e_1$ 和 $e_2$ 后	<b>0.79</b>	<b>0.89</b>	<b>907</b>
$e_2$ 和 $e_3$ 后	0.82	0.91	887
$e_1$ 和 $e_3$ 后	0.84	0.92	882
$e_1, e_2$ 和 $e_3$ 后	0.80	0.89	911

4.3 SADAЕ 鲁棒性比较

火焰图像质量易受成像系统和环境干扰,其基本特征对高精度燃烧状态预测至关重要<sup>[18-19]</sup>。本文通过对比实验验证 SADAЕ 的特征学习能力。如图 11 所示,不同噪声水平下,各模型表现差异显著。当噪声  $\delta$  增加时, pix2pix<sup>[20]</sup> 模型精度下降,表明其抗噪能力弱,难以提取噪声图像中的有效信息。相比之下, CAE<sup>[12]</sup>、AST-GCN<sup>[5]</sup>、ViT<sup>[16]</sup> 和 MSViL 模型在噪声初期性能未显著下降,甚至略有提升,随后随噪声增加逐渐变差,表明其在 SADAЕ 作用下表现出较强抗噪能力,整体性能显著提升。其中, SADAЕ 与 ViL 模型结合表现最佳, RMSE=3.00,验证了其鲁棒性和优越性。

为了证明 SADAЕ 的出色特征学习能力,本文将其与该领域其他先进特征学习方法进行了比较,包括 CE<sup>[21]</sup>、CBAM<sup>[22]</sup>、CDE-SSA<sup>[23]</sup> 和 ADAЕ<sup>[14]</sup>。ADAЕ 与 SADAЕ 结构相似,噪声水平均设为 0.1 以确保公平性。由于成像系统和环境干扰会影响火焰图像质量,预测方法的鲁棒性至关重要。因此,在测试图像中添加了不同级别的高斯噪声,预测精度如图 12 所示。随着噪声水平增加,预测性能逐渐下降,主要因低质量图像包含大量无效信息。当噪声  $\delta$  增至 0.1 时, CE-ViL、CBAM-ViL 和 CDE-SSA 的预测精度已经开始大幅降低。当  $\delta$  增至 0.2 时, ADAЕ-MSViL

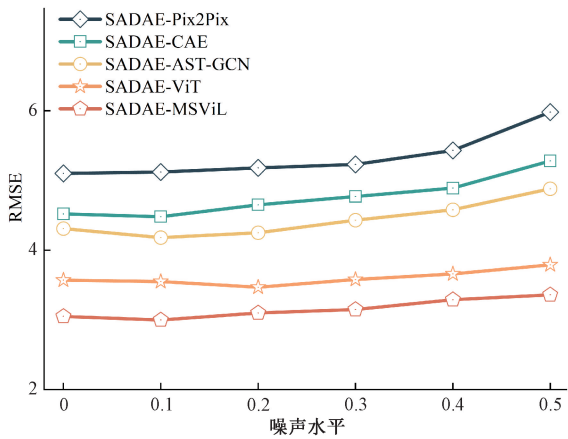


图 11 不同噪声水平下各模型的特征提取比较  
Fig. 11 Comparison of feature extraction of each model under different noise levels

模型性能迅速恶化,表明其抗干扰能力有限。然而,SADAe-MSViL 性能变化较小,即使噪声水平达到 0.5,其精度仍保持在 0.9 以上。实验结果证实,SADAe 能够有效提取火焰图像的鲁棒特征,展现出优异的抗噪能力。

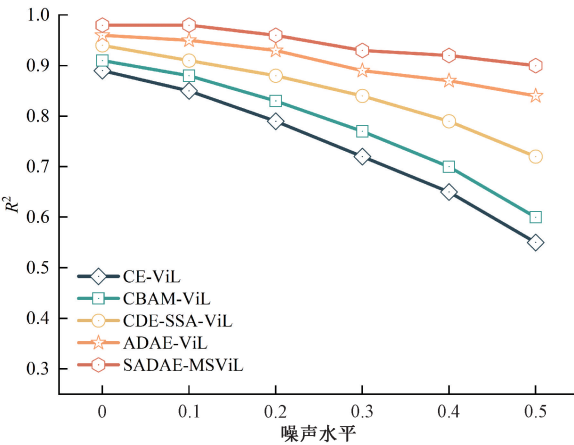


图 12 基于特征提取的模型的鲁棒性测试

Fig. 12 Robustness test of the model based on feature extraction

4.4 ViL 门控机制与参数选择

为了验证 ViL 模型中门控非线性低秩注意力机制相比传统注意力机制的优势,本文分别对门控模块和低秩维度的设计进行了对比实验研究。

由表 4 可知,随着低秩维度的增加,模型整体性能呈现出逐步提升的趋势,其中非线性门控机制在各个维度下均优于原始 Linformer 机制,尤其在 128 维时表现最优,且实现了性能与效率之间的良好平衡。相比之下,原始 Linformer 在相同维度下误差更大,说明引入的非线性门控机制在提高预测精度的同时大幅减少运行时间。因为大小为  $16 \times 16$  时图像块可以分割为 256 个,降维后的维度应小于原始序列长度,即不超过 256,否则无法实现降维效果。

表 4 ViL 非线性门控机制对模型的影响

Table 4 The impact of ViL nonlinear gating mechanism on the model

类型	低秩维度	MAE	RMSE	Time/min
非线性门控	64	2.45	3.33	106
	128	2.22	3.00	118
	256	2.20	2.98	131
原始 Linformer	64	2.70	3.52	92
	128	2.48	3.18	108
	256	2.45	3.15	125

Linformer 的数量和 MHSA 的数量都会影响模型效果。本文通过实验比对选择了最佳参数。适当增加 Linformer 的数量  $\beta$  可以增强模型对特征的捕捉能力。从图 13 可以看出,随着  $\beta$  的增加,模型的预测性能逐步提升,同时训练时间也相应增长。本文选择  $\beta=12$ ,此时虽然训练时间相比  $\beta=8$  时增加 23 min,但 RMSE 和 MAE 分别降低了 7.5% 和 13.6%,体现了较高的性价比。

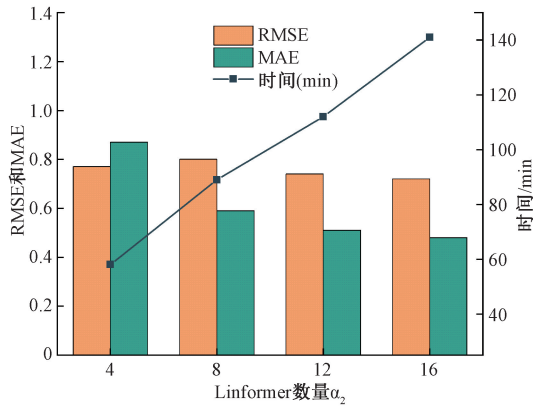


图 13 Linformer 数量比较

Fig. 13 Linformer quantity comparison

更多的 MHSA 数量能够提升模型对不同特征维度的捕捉效果,但过多的 MHSA 数会显著增加参数量和计算开销,因此需要在精度与效率间找到平衡点。由图 14 可知,当 MHSA 数量为 12 时, RMSE 和 MAE 达到最小的同时,计算时间仅为 112 min。

4.5 多尺度 ViL 比较

本文通过对比实验验证了多尺度融合机制的优势。由表 5 可知,单一尺度模型训练时间短,但精度较低,无法同时捕捉局部细节与全局信息。8+16 组合在 MAE 和 RMSE 上表现最优,优于其他多尺度组合。尝试 8+16+32 三尺度融合后, MAE 和 RMSE 略有提升,但训练时间显著增加。本文最终选择 8+16 作为最优组合。结果表明,多尺度融合能有效兼顾高频细节与低频全局特征,8+16 组合在性能与效率间取得了最佳平衡。

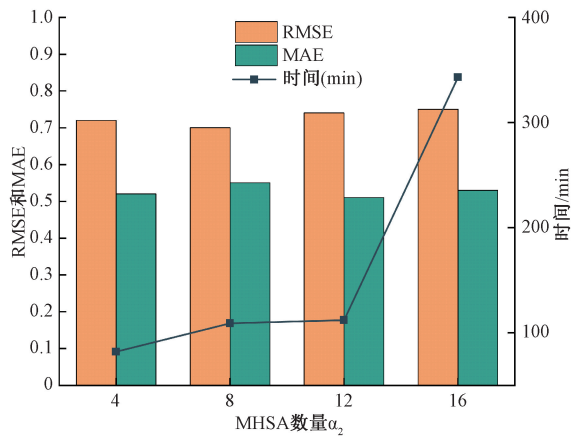


图 14 多头自注意力数量比较

Fig. 14 Comparison of the number of self-attention heads

表 5 不同尺度图像块对模型的影响

Table 5 The impact of image blocks of different scales on the model

尺度	MAE	RMSE	Time/min
8	2.56	3.50	89
16	2.68	3.62	95
32	2.97	3.88	98
8+16	<b>2.22</b>	<b>3.00</b>	<b>118</b>
8+32	2.29	3.23	108
16+32	2.35	3.48	102
8+16+32	2.18	2.98	132

4.6 模型比较

为了探究所提出模型的优势, 本文将其与当前几种先进模型 (包括 PPIN<sup>[24]</sup>、MJDNN<sup>[25]</sup>、ADAE-LSSVM<sup>[14]</sup> 和 MSViL) 进行了全面对比, 如图 15 所示。实验结果表明, SADAE-MSViL 在模型拟合效果上均显著优于其他对比模型。表明本文提出的模型在捕捉复杂数据特征和提升预测精度方面具有显著优势。

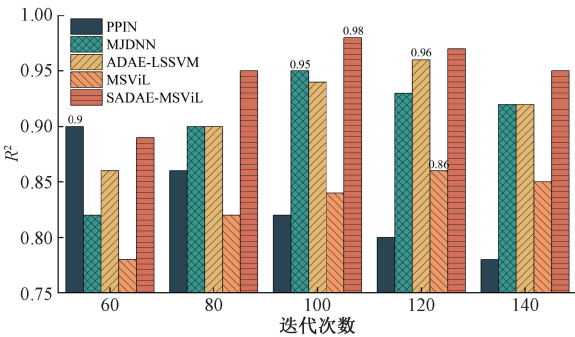


图 15 模型比较

Fig. 15 Comparison of advanced models

5 结 论

本文提出了一种 SADAE-MSViL 模型, 利用生物质燃料燃烧产生的火焰温度图像, 实现对 NO<sub>x</sub> 排放量的预测。该模型充分提取火焰深层特征并进行特征预测。提出 SADAE 模型, 利用 SA 模块从高维火焰图像中提取深度特征, 确定最优噪声水平  $\delta = 0.1$ , 优于其他先进特征提取模型; 设计 MSViL 模型, 通过低秩门控机制将高维注意力矩阵降维为低秩形式, 降低计算维度并实现精度和运行时间之间的平衡, 确定 Linformer 和 MHSA 最优数量为 8 和 12; 提出的 SADAE-MSViL 模型有较高的 NO<sub>x</sub> 预测精度, RMSE = 3.00,  $R^2 = 0.98$ , 优于 PPIN、MJDNN、ADAE-LSSVM 和 ViL 模型并展现出较高的准确性和可靠性。未来的工作将针对不同燃烧设施, 定制并优化所提出的 NO<sub>x</sub> 排放预测方法。

参考文献

[1] 殷柳玲, 曹希, 卜昌盛. 煤富氧燃烧及污染物排放特性研究进展[J]. 电力科技与环保, 2025, 41(1): 96-109.  
YIN L L, CAO X, BU CH SH, et al. Research progress on coal oxy-fuel combustion and pollutant emission characteristics[J]. Electric Power Technology and Environmental Protection, 2025, 41(1): 96-109.

[2] XU Q Q, HAO X CH, SHI X, et al. Control of denitration system in cement calcination process: A novel method of deep neural network model predictive control[J]. Journal of Cleaner Production, 2022, 332.

[3] TANG ZH H, SUI M, WANG X, et al. Theory-guided deep neural network for boiler 3-D NO<sub>x</sub> concentration distribution prediction [J]. Energy, 2024, 299: 131500.

[4] 刘建军, 赵旭, 张卫东, 等. 基于 CNN(1D)-LSTM 模型的电站锅炉 SCR 入口 NO<sub>x</sub> 浓度预测[J]. 电子测量技术, 2024, 46(13): 59-65.  
LIU J J, ZHAO X, ZHANG W D, et al. Prediction of SCR inlet NO<sub>x</sub> concentration in power plant boilers based on CNN (1D)-LSTM model [J]. Electronic Measurement Technology, 2024, 46(13): 59-65.

[5] 周永清, 郝大伟, 樊昱晨, 等. 基于时空注意力图卷积网络的锅炉 NO<sub>x</sub> 预测方法[J]. 煤炭学报, 2024, 49(10): 4127-4137.  
ZHOU Y Q, HAO D W, FAN Y CH, et al. A boiler NO<sub>x</sub> prediction method based on spatio-temporal attention graph convolutional network[J]. Journal of China Coal Society, 2024, 49(10): 4127-4137.

[6] 邢红涛, 郭江龙, 刘书安, 等. 基于 CNN-LSTM 混合神经网络模型的 NO<sub>x</sub> 排放预测[J]. 电子测量技术, 2024, 45(2): 98-103.  
XING H T, GUO J L, LIU SH AN, et al. NO<sub>x</sub> emission prediction based on CNN-LSTM hybrid neural network model [J]. Electronic Measurement Technology, 2024, 45(2): 98-103.

[7] LIU Y, ZHOU J, FAN W. A novel robust dynamic



- method for NO<sub>x</sub> emissions prediction in a thermal power plant[J]. The Canadian Journal of Chemical Engineering, 2023, 101(5): 2391-2402.
- [8] 周榕, 王博强, 王锴. 基于卷积神经网络的火焰图像识别技术的研究[J]. 热能动力工程, 2024, 39(2): 184-193.
- ZHOU R, WANG B Q, WANG K. Research on flame image recognition technology based on convolutional neural network[J]. Journal of Engineering for Thermal Energy and Power, 2024, 39(2): 184-193.
- [9] LI N, LU G, LI X L, et al. Prediction of NO<sub>x</sub> emissions from a biomass fired combustion process based on flame radical imaging and deep learning techniques[J]. Combustion Science and Technology, 2016, 188(2): 233-246.
- [10] XAYASOUK T, LEE H M, LEE G. Air pollution prediction using long short-term memory(LSTM) and deep autoencoder(DAE) models[J]. Sustainability, 2020, 12(6): 2570.
- [11] TANG ZH H, WANG SH K, CHAI X Y, et al. Auto-encoder-extreme learning machine model for boiler NO<sub>x</sub> emission concentration prediction[J]. Energy, 2022, 256: 124552.
- [12] 王宝华, 温武斌, 王建超, 等. 基于火焰图像和半监督学习模型的燃烧 NO<sub>x</sub> 排放预测[J]. 节能技术, 2024, 42(6): 517-522.
- WANG B H, WEN W B, WANG J CH, et al. Prediction of combustion NO<sub>x</sub> emissions based on flame images and semi-supervised learning model[J]. Energy Conservation Technology, 2024, 42(6): 517-522.
- [13] QIN L, LU G, HOSSAIN M M, et al. A flame imaging-based online deep learning model for predicting NO<sub>x</sub> emissions from an oxy-biomass combustion process[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 71: 1-11.
- [14] HAN ZH ZH, XIE Y, HOSSAIN M M, et al. A hybrid deep neural network model for NO<sub>x</sub> emission prediction of heavy oil-fired boiler flames[J]. Fuel, 2023, 333: 126419.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [16] 程思雨, 陈莹. 基于 ViT 的细粒度特征增强无监督行人重识别方法[J]. 电子测量与仪器学报, 2024, 38(9): 24-35.
- CHENG S Y, CHEN Y. Fine-grained feature enhancement for unsupervised person re-identification based on ViT[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 24-35.
- [17] 代正华, 孟凯, 郭庆华, 等. 基于双色法油浆火焰温度场和碳烟浓度场分析[J]. 华东理工大学学报(自然科学版), 2020, 46(3): 319-325.
- DAI ZH H, MENG K, GUO Q H, et al. Analysis of temperature field and soot concentration field in slurry flame based on two-color method[J]. Journal of East China University of Science and Technology (Natural Science Edition), 2020, 46(3): 319-325.
- [18] 梁秀满, 安金铭, 曹晓华, 等. 基于改进 MobileNetV3 烧结断面火焰图像识别[J]. 电子测量技术, 2023, 46(14): 182-187.
- LIANG X M, AN J M, CAO X H, et al. Sintering section flame image recognition based on improved MobileNetV3[J]. Electronic Measurement Technology, 2023, 46(14): 182-187.
- [19] 王安, 温武斌, 刘宏文, 等. 基于卷积自编码的火焰图像稳定性定量评估[J]. 动力工程学报, 2025, 45(3): 325-333.
- WANG AN, WEN W B, LIU H W, et al. Quantitative evaluation of flame image stability based on convolutional autoencoder[J]. Journal of Power Engineering, 2025, 45(3): 325-333.
- [20] 张志宇, 李新利, 杨国田. 基于改进生成对抗网络 pix2pix 的火电厂 NO<sub>x</sub> 排放预测[J]. 中国电机工程学报, 2023, 43(12): 4678-4688.
- ZHANG ZH Y, LI X L, YANG G T. NO<sub>x</sub> emission prediction in thermal power plants based on improved pix2pix generative adversarial network[J]. Proceedings of the CSEE, 2023, 43(12): 4678-4688.
- [21] WANG ZH, ZHOU H CH, PENG X Y, et al. A predictive model with time-varying delays employing channel equalization convolutional neural network for NO<sub>x</sub> emissions in flexible power generation[J]. Energy, 2024, 306: 132495.
- [22] LYU Y, QI X Y, ZHENG X, et al. Unsupervised quantitative judgment of furnace combustion state with CBAM-SCAE-based flame feature extraction[J]. Journal of the Energy Institute, 2024, 116: 101733-101733.
- [23] 宋美艳, 刘畅, 张津, 等. 基于麻雀算法优化支持向量机的 NO<sub>x</sub> 浓度预测[J]. 计算机仿真, 2024, 41(7): 129-134.
- SONG M Y, LIU CH, ZHANG J, et al. NO<sub>x</sub> concentration prediction based on sparrow algorithm optimized support vector machine[J]. Computer Simulation, 2024, 41(7): 129-134.
- [24] 任少君, 朱保宇, 翁琪航, 等. 基于物理信息神经网络的燃煤锅炉 NO<sub>x</sub> 排放浓度预测方法[J]. 中国电机工程学报, 2024, 44(20): 8157-8166.
- REN SH J, ZHU B Y, WENG Q H, et al. NO<sub>x</sub> emission concentration prediction method for coal-fired boilers based on physics-informed neural network[J]. Proceedings of the CSEE, 2024, 44(20): 8157-8166.
- [25] TANG ZH H, WANG SH K, LI Y. Dynamic NO<sub>x</sub> emission concentration prediction based on the combined feature selection algorithm and deep neural network[J]. Energy, 2024, 292: 130608.

## 作者简介

常胜君, 硕士研究生, 主要研究方向为生物质燃料 NO<sub>x</sub> 排放预测、火焰图像处理。

E-mail: 1208247159@qq.com

郝润芳(通信作者), 副教授, 主要研究方向为人工智能、信号与信息处理。

E-mail: haorunfang@tyut.edu.cn