

文章编号: 1006-3080(2025)03-0380-11

DOI: 10.14135/j.cnki.1006-3080.20240614001

基于膨胀因果卷积和注意力机制的气体识别方法

俞凌伟, 杨孟平, 杨海, 王喆

(华东理工大学信息科学与工程学院, 上海 200237)

摘要:提出了一种基于膨胀因果卷积和注意力机制的气体识别方法, 该算法结合 Transformer 中的注意力机制和多尺度时态卷积网络提取全局和局部特征, 获得了更具表示性的特征和更大的感受野, 捕获气体的瞬时信息和变化趋势。在 Open Sampling、Drift、Twin 3 个不同数据集上进行了实验, 结果表明, 所提出的方法分别达到 99.47%、99.61% 和 99.22% 的准确率, 优于现有主流方法, 证实了其有效性。

关键词:电子鼻系统; 气体识别; 注意力机制; 时态卷积网络; 时间序列

中图分类号: TP391

文献标志码: A

嗅觉传感是通过模仿人类嗅觉系统的原理和功能, 利用传感技术实现对气味、气体和化学物质的检测与识别。作为嗅觉传感的核心部件, 电子鼻系统采用交叉敏感传感器阵列来获取气体信息。由于其客观和快速的检测能力, 电子鼻系统在食品工程、化学检验和医疗卫生中得到了广泛应用^[1-4]。然而, 有两个问题限制了它在气体识别中的广泛应用, 一是半导体气体传感器在长期工作过程中会发生不规则漂移, 这是由老化、温度、空气流速和环境造成的^[5]; 二是大多数半导体气体传感器的选择性较差, 只对特定气体反应明显, 很难在复杂的大气中准确地检测目标气体^[6]。为了提高气体检测的选择性并消除不规则漂移的影响, 可以使用机器学习技术融合来自多个气体传感器的信号。因此, 本文从特征提取的角度出发, 提出一种基于膨胀因果卷积和注意力机制的气体识别方法, 将其与气体检测相结合, 学习短期传感器响应并消除传感器漂移的影响, 从而准确识别目标气体。

在传统的气体识别中, 常用的典型算法包括 K-最近邻 (KNN)^[7]、决策树 (DT)^[8] 和随机森林 (RF)^[9]。在研究人员构建特殊的人工特征后, 大多数传统的

机器学习方法都可以很好地工作^[10-11]。然而当发生严重漂移时, 这些显式特征很难区分信号所代表的气体, 从而导致传统机器学习方法的失败。近些年的研究经常利用卷积神经网络 (CNN)^[12-15] 等深度学习方法。在 CNN 架构中, 卷积层负责特征提取和处理, 而全连接层负责分类。通过集成特征提取、处理和模式识别, CNN 最大限度地减少了气体识别中的人工干预^[16]。然而, CNN 在气体识别中的应用仍存在一些问題。首先, 对捕获数据中的依赖性起决定性作用的 CNN 感受域, 仅限于局部区域, 这阻碍了其有效捕获气体特征中的长期依赖性。其次, 从电子鼻系统获得的气体信息由不同时间点的各种传感器检测数据组成, 不同检测时间下的气体信息表达了不同物理含义。对于气体识别任务, CNN 并不能很好地挖掘数据之间的全局联系。

为了解决感受域有限的问题, 本文使用膨胀因果卷积代替传统卷积操作, 膨胀因果卷积中每个像素点只依赖于其之前的输入像素点, 而不受后续像素点的影响, 允许网络考虑到数据中的时序依赖关系^[17]。此外, 通过增加卷积核的空间感受野, 使得每个输出像素能够感知到更广泛的输入信息, 从而提

收稿日期: 2024-06-14

基金项目: 国家重点研发计划项目(2022YFB3203500)

作者简介: 俞凌伟(2000—), 男, 浙江杭州人, 硕士生, 主要研究方向为气体识别。E-mail: 517690329@qq.com

通信联系人: 王喆, E-mail: wangzhe@ecust.edu.cn

引用本文: 俞凌伟, 杨孟平, 杨海, 等. 基于膨胀因果卷积和注意力机制的气体识别方法 [J]. 华东理工大学学报(自然科学版), 2025, 51(3): 380-390.

Citation: YU Lingwei, YANG Mengping, YANG Hai, *et al.* Gas Recognition Method Based on Dilatation Causal Convolution and Attention Mechanism[J]. Journal of East China University of Science and Technology, 2025, 51(3): 380-390.

高了模型的感知范围。Transformer 利用多头注意力机制来提取特征并建立相关性^[18],其稳健的全局建模能力已被广泛应用于图像识别^[19]、超分辨率^[20]等领域。为了提升网络的全局信息建模能力,获得更好的局部最优解,本文使用 TE(Transformer encoder)模块,将传感器的每个时间特征视为句子中的一个词向量,来挖掘传感器信号对之间潜在的关系,以获得全局特征。

本文提出了一种基于注意力机制的多尺度时态卷积神经网络(Multiscale Temporal Convolutional Neural Network with Attention, MTCNA),它由3个关键部分组成:第一为多尺度时态卷积网络,将原始数据使用多尺度学习策略进行扩充,增强数据时间维度上的多样性,然后对每个时间尺度的数据都使用独立的时态卷积网络提取局部特征和时间信息;第二为 TE 模块,用于捕捉全局特征;第三为结合中心损失的交叉熵损失函数,交叉熵损失用于确保模型分类准确性,而中心损失^[21]用于提高特征的判别性,两种损失函数结合关注类间距离和类内距离,可以提高模型的泛化能力。

1 相关工作

1.1 电子鼻系统中气体数据的特征

电子鼻系统具备快速、可重复的气体检测能力,并通过交叉敏感传感器阵列获取气体指纹信息。交叉灵敏度意味着同一个传感器对不同的气体可能产生不同的响应;同样,同一种气体也可能引发不同传感器的响应。因此,单个传感器的响应曲线携带了关于其他气体的丰富信息,不同传感器根据其灵敏度和选择性也会表现出不同程度的响应^[22-23]。

电子鼻系统采用的标准检测方法是动态顶空法,该方法生成的传感器响应曲线呈现上升和下降两个不同的阶段。上升阶段的发生是由于高顶空空气富集迅速进入传感器气室。相反,下降阶段是由样品气体挥发和洁净空气之间的动态平衡产生的。这些不同的响应代表了不同的气体挥发状态。受采样频率和气体浓度的影响,上升相位响应变化迅速,包含高频气体信息,而下降相位响应变化较慢,包含低频气体信息。由于这些阶段包含不同的信息,因此这些阶段的响应对决策绩效具有不同的含义,不同响应之间存在潜在的相关性,可以利用这些相关性来提高模型性能。

1.2 气体识别研究现状

传统的气体识别方法是从原始检测信号中提取

关键特征,以减轻冗余信息对分类性能的影响。然后,进行特征处理降低特征之间的相关性和非线性,从而得到最优特征集。最后,模式识别技术被应用于这个特征集,成功识别气体。然而,该方法数据处理的复杂性引入了过多的人工干预,导致分类模型的稳定性较差。

CNN 是一种经典的深度学习算法,由于其强大的特征提取能力,CNN 已被证明在气体识别方面非常有效。Guo 等^[24]将 CNN 与 LSTM(Long-Short-Term Memory)网络相结合,从气体信息中全面提取时间特征。通过有效捕获随时间变化的模式,可以准确地识别气体样本。然而,LSTM-CNN 混合算法的计算复杂度和训练时间可能相对较高,限制了其实时性。Xiong 等^[25]提出了一种将峰值神经网络(SNN)与残差网络(ResNet)相结合的新方法来识别变质的食物气体。SNN 和 ResNet 的集成使其能够处理基于峰值的数据,并通过剩余连接保留关键信息,在识别具有复杂时间动态的气体方面表现出优异的性能。Ma 等^[26]将气体传感器阵列的多维动态响应信号作为图像处理,将动态响应图与 CNN 相结合,提高了模型性能。这种基于图像表示增强了对气体特征的空间上下文理解,从而实现了有效的特征提取。尽管如此,这种方法严重依赖于将传感器数据转换为类似图像的输入,可能会丢失细粒度的信息。Lee 等^[27]利用基于 CNN 的深度学习技术,设计了一种用于医疗保健的便携式电子鼻系统。他们的方法允许呼吸系统疾病的诊断,展示了 CNN 在各种应用场景中的多功能性。然而,传统的 CNN 架构具有有限的感受野,这阻碍了它们捕获气体特征的长程相关性的能力,这可能导致模型性能的损失。为了解决这一限制,本文提出的 MTCNA 使用膨胀因果卷积来代替普通卷积,有效解决了 CNN 感受野有限的问题。

Transformer 已经在各个领域展示了其多功能性,包括语言处理、图像识别、对象检测、图像超分辨率等。研究人员将 Transformer 的优势与 CNN 的优势结合起来,以进一步增强其功能。Srinivas 等^[28]将自注意机制集成到 ResNet 架构中,取代了传统的卷积,在不显著改变网络结构的情况下提高了模型性能。使用自注意力机制的优势在于它能够捕获数据中的长期依赖关系,缺乏显式的局部信息提取可能会妨碍其在需要精确空间关系任务上的性能。Cordonnier 等^[29]证明,自注意机制可以通过调整特定参数来近似卷积层的效果,这种方法受益于自注意力机制提供的全局上下文理解和空间信息处理。然

而,该方法的性能可能对参数设置很敏感。Yuan 等^[30]在补丁切片前应用了卷积和最大池化,使模型能够高效地捕获图像的主要特征。Xiao 等^[31]提出在初始特征提取阶段加入多重卷积来增强模型泛化。然而,电子鼻系统具有很少的检测数据,而以上提出的模型参数众多,这导致难以有效地优化参数。通过与不同的模块有效协作,MTCNA 解决了以前的方法需要大量样本的问题,这使得它更适合在这种情况下处理气体信息。

2 实验部分

2.1 整体网络结构

本文提出的 MTCNA 的结构示意图如图 1 所示,网络整体结构由多尺度时态卷积网络(TCN)、TE 模块和全连接层组成。首先对输入数据采用多尺度学习策略,生成不同时间尺度的信号,产生更全面、更广泛的信号理解表示,增强模型的特征提取能力和识别性能。不同时间尺度的数据分别进入各自对应的时态卷积子网络,通过膨胀因果卷积,网络可随着层深度增加而增大感知域,对气体传感器数据中的局部特征进行多尺度提取。其次,TE 模块利用多头注意力机制来捕捉不同传感器数据在不同时间尺度上特征之间的全局相关性,放大不同时间尺度特征之间的异同,生成新的特征表示。最后,新的特

征经过全连接层后,模型输出所有类别的预测概率。此外,模型在使用交叉熵损失(L_{CE})的基础上,还使用中心损失(L_{Center})作为辅助项以最小化类内损失。

2.2 多尺度时态卷积网络

2.2.1 多尺度学习策略 多尺度学习策略在不同尺度下提取不同的时间细节,当输入信号包含异常响应时,网络识别能力更强。当传感器以低采样率运行时,其信号会体现出长期的响应变化。而当传感器以高采样率运行时,则可能同时显示出长期和短期的响应变化。长期响应和短期响应都有助于气体分类,但长期响应更为重要,这是因为它反映的是气体浓度的变化,而短时响应包括空气流量、噪声、传感材料变化等次要信息。在气体传感器信号中存在不稳定响应的情况下,多尺度学习策略在提取时间模式时表现更好,因为它增强了长期响应的贡献。

本文采用一种降采样方法来生成不同尺度的信号。假设有一个气体传感器信号 $\{x_i\}=\{x_1, x_2, x_3, \dots, x_N\}$,其中 x_i 表示气体传感器在时刻 i 的电压值,每个信号有 N 个时刻的数据。受 MCNA(Multiscale convolutional neural network with attention)的启发,本文用降采样率 2^k ($k=0,1,2,\dots$)在不同的尺度上构建信号图 y_i 。通过获得原始信号中每 2^k 个数据的最大值而不是平均值来进行降采样过程,如图 2 所示。使用式(1)计算尺度 k 处的信号图。

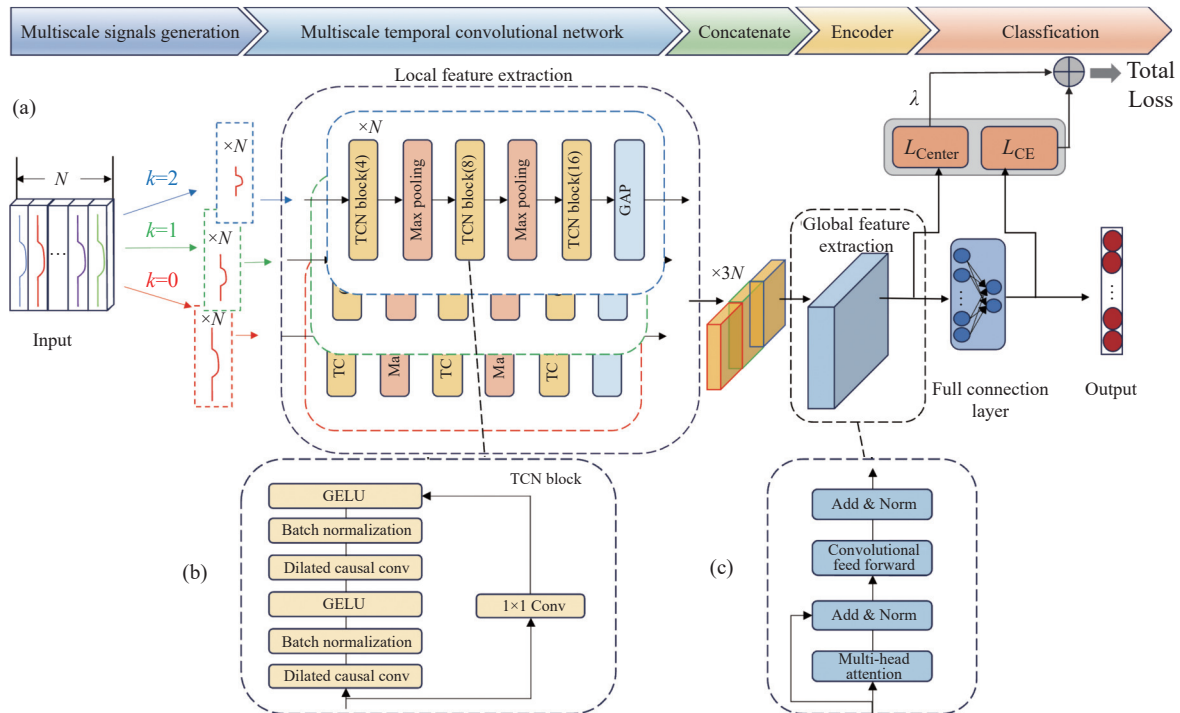


图 1 融合注意力机制的多尺度时态卷积神经网络结构图 (a); TCN 结构图 (b); TE 模块结构图 (c)

Fig. 1 Multi scale temporal convolutional neural network structure diagram of fused attention mechanism (a); TCN structure diagram (b); Transformer encoder structure diagram (c)

$$y_j^k = \max_{(j-1)2^k \leq i \leq j \cdot 2^k} \{x_i\}, 1 \leq j \leq \text{Round} \frac{N}{2^k} \quad (1)$$

其中 $\text{Round}(\frac{N}{2^k})$ 表示取整到小于或等于 $(\frac{N}{2^k})$ 的最接近整数。当 k 取值较小时, 得到的是低尺度数据; 当 k 取值较大时, 得到的是高尺度数据。从感受野的角度来看, 低尺度数据的感受野较小, 更关注局部信息; 高尺度数据的感受野较大, 更关注全局信息。从语义信息角度来看, 低尺度数据关注底层的特征信息, 高尺度数据关注更高层次的语义信息。这些不同层次的信息之间存在依赖关系, 高层次语义需要依托于底层特征, 因此不同尺度的数据也存在内在联系。而本文使用的多尺度学习策略就是将低尺度和高尺度数据结合起来, 通过共享参数的方式进行学习, 有助于模型更全面地捕获数据的语义信息, 从而提高气体识别的性能和鲁棒性。

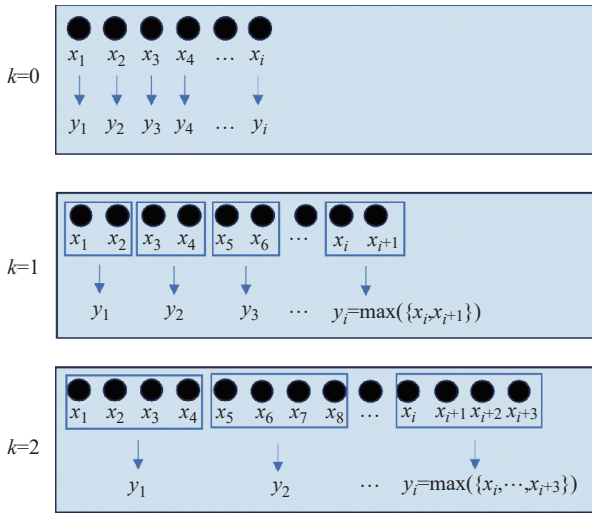


图 2 不同采样率下信号的示意图
Fig. 2 Schematic diagram of signal at different sampling rates

2.2.2 时态卷积子网络 MTCNA 的时态卷积子网络由多个并行 TCN 块组成, 用于提取当前时间尺度的气体数据的局部特征。传统卷积在处理长序列时存在信息损失或者梯度消失的问题, 而 TCN 具有捕获长时间序列依赖关系的能力, 能适应不同时间尺度的特征。本文的 TCN 块基于 Bai 等^[32] 提出的架构, 详细的架构如图 1(b) 所示。

TCN 块主要由膨胀因果卷积组成。因果卷积使得时刻 t 的输出只与时刻 t 以及之前的输入相关, 如图 3(a) 所示, 其中零填充用于产生与输入相同长度的输出序列。

因果卷积的引入允许神经网络考虑气体数据中存在的顺序依赖性, 即时间上的因果关联。然而, 与传统 CNN 类似, 因果卷积也受到卷积核 (k') 大小的限制, 这限制了输出元素所能考虑的时间步长。为

了捕获更长的依赖关系, 通常需要构建更深的网络。因此, 引入了膨胀因子 (d) 的概念, 以在保持网络较浅的同时, 增加输出部分对当前时刻之前更多历史数据的利用。膨胀因子代表了输入序列元素之间的距离, 这些元素用于计算输出结果。

由于输出值的输入项相邻, 因此, 传统的卷积层可以被认为是一个膨胀因子为 1 的扩展卷积层, 在分析时间序列时, 卷积核大小和膨胀因子可以分别看作是采样次数和采样周期。对于相同的采样次数, 采样周期越长, 包含的时间信息越多。使用膨胀因子会产生更大的时间步长, 同时避免与下采样相关的信息损失, 因为与下采样不同, 它不会修改输出特征的长度。如图 3(b) 所示, 膨胀卷积通过间隔采样扩展感受域, 使网络回到 $1 + \sum_{i=0}^{n-1} (k-1)d_i$ 时间步长, 其中 n 为层数。可以看出, 在层数相同的情况下, 膨胀因果卷积的输出比不含膨胀因子的因果卷积获得了更多的时间序列信息。扩展的因果卷积允许网络像 RNN 一样依赖于长期的历史信息, 同时保留了 CNN 的并行性和多尺度特征提取能力。

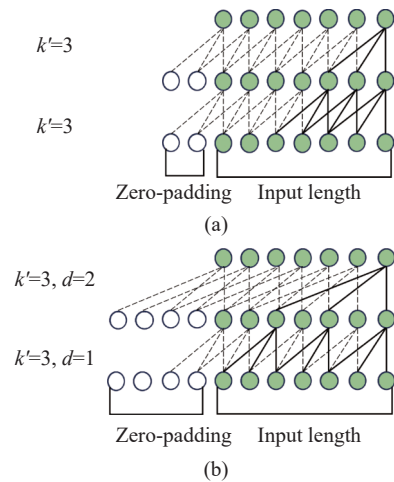


图 3 不含膨胀因子的因果卷积 (a); 膨胀因果卷积 (b)
Fig. 3 Causal convolution without expansion factor (a) and with expansion factor (b)

此外, TCN 单元还包含残差连接, 这使得该结构能够获得稳定的深度网络。残差块由两个卷积层和非线性激活函数 GELU 组成, 每一层加入批量归一化, 使网络正则化, 增强其泛化能力。

2.3 TE 模块

本文使用的 TE 模块的结构如图 1(c) 所示, 主要由多头注意力和卷积前馈层组成。本文对原始的前馈层进行了改进, 增加了一个卷积层, 增强了网络对时间序列数据的绝对位置信息进行编码的能力。多头注意力机制的具体结构如图 4 所示, 利用多个平行的自注意力层从多个方面获取全局信息, 其方程如

式(2)和式(3)所示:

$$\text{Multihead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{concat}(\text{head1}, \dots, \text{headH})\mathbf{W}^0 \quad (2)$$

$$\text{head1} = \text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d\mathbf{K}}}\right)\mathbf{V} \quad (3)$$

其中, \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 分别表示 Query、Key、Value 的矩阵, 它们是由经过不同线性变换的输入特征向量构成的矩阵, \mathbf{W} 为权重矩阵。将它们视为输入特征向量在不同空间上的投影, 具体的线性变换矩阵参数通过学习得到。将 Query 矩阵与 Key 矩阵相乘得到当前传感器的信号特征与其他传感器的信号特征间的相似度。与此类似, Value 矩阵与相似特征矩阵相乘可以在原始特征的基础上更清晰地关注重要特征。然而, 在对当前位置的信息进行编码时, 自注意力机制会过于关注其位置, 而忽略其他位置。因此, 引入具有更大感受野的多头注意力机制, 有效地对不同特征之间的相关性进行建模。多头注意力机制通过产生 H 组不同线性投影来对 \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 进行变换, 其中 H 为注意头的数量。这为分配注意力权重提供了更多选择, 并增加了公平分配的可能性。此外, 多头注意力机制还使注意层的输出包含来自不同子空间的编码表征信息, 从而提高模型的表示能力。

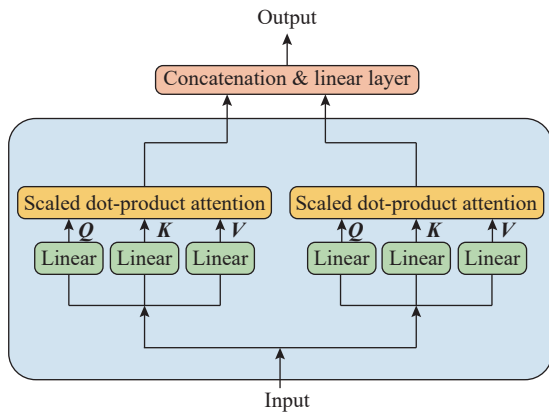


图 4 多头注意力机制的详细结构图

Fig. 4 Detailed structure of the multihead attention mechanism

此外, 注意力机制是置换不变性的, 即不能捕获特征之间的位置关系。本文在 Transformer 中使用卷积前馈网络代替原来的前馈网络。图 5 所示为改进前后的前馈网络的对比结果, 其中图 5(b) 所示为卷积前馈模块的具体结构, 卷积前馈网络在第一个全连接层(FCL)和激活函数之间增加了一个一维深度卷积(DWConv), 将整流线性单元(ReLU)激活函数改为高斯误差线性单元(GELU)激活函数。深度卷积由可分离卷积和逐点卷积组成, 该结构与传统卷积相当, 可以应用于特征提取, 但它具有更少的参数和更低的操作成本。零填充的一维深度卷积捕捉相

邻 token 之间的关系, 获得特征的绝对位置信息, 保持网络结构的平移不变性, 使输入序列排列可变但平移不变, 加强局部上下文之间的联系。因此, TE 模块增强了网络的鲁棒性, 减少了异常值带来的影响, 特别是当气体成分突然波动时, 识别更加准确。

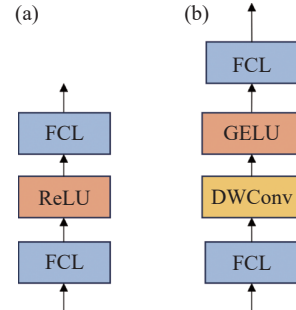


图 5 原始前馈 (a); 卷积前馈 (b)

Fig. 5 Original feedforward (a); convolutional feedforward (b)

2.4 损失函数

在经典的分类任务中, 损失函数通常使用交叉熵损失。然而, 传统的交叉熵损失往往只关注类别信息, 而忽略了特征之间的空间关系, 导致模型学到的特征不够紧凑和区分度不够。为了在模型学习过程中对特征更好地聚类, 本文引入中心损失来让同一类别的样本在特征空间中更加接近, 通过最小化样本特征与其对应类别中心的距离, 中心损失能够提高特征的紧凑性, 从而增强模型的判别能力。中心损失的计算公式如式(4)所示:

$$L_{\text{Center}} = \frac{1}{2} \sum_{i=1}^N D(f(x^i), c_{y_i}) \quad (4)$$

其中, D 表示欧几里得距离的平方, c_{y_i} 是类 y_i 的中心, f 表示模型输出的特征, 对应图 1(a) 中 TE 模块和全连接层之间的特征。总体损失公式如式(5)所示:

$$L_{\text{Total}} = L_{\text{CE}} + \lambda L_{\text{Center}} \quad (5)$$

其中 L_{CE} 、 L_{Center} 分别代表气体分类任务中的交叉熵损失和中心损失, 利用参数 λ 来控制交叉熵损失和中心损失之间的比例, 模型最终用 L_{Total} 作为总损失函数来训练模型。

3 实验结果与分析

3.1 数据集介绍

本文使用 3 个电子鼻数据集来验证所提方法的有效性, 每个数据集是在不同检测环境中使用不同的电子鼻传感器采集。

Open Sampling 数据集是从传感器阵列模块中收集, 模块有 8 个金属氧化物(MOX)气体传感器, 其在复杂的环境条件下检测和识别潜在危险的化学气体

物质。传感器阵列模块的响应信号是从集成传感器阵列的 72 个气体传感器中,以每个气体传感器的有源敏感膜上的电阻形式读出。每次测量都会产生一个 72 通道的时间序列,每个通道都由一个 260 s 的时间序列表示,采样率 100 Hz,反映了评估场景中的所有环境变化。该数据集的代表性样本如图 6 所示。该数据集包含 10 类不同纯化学气体的记录,包括丙酮、乙醛、氨、丁醇、乙烯、甲烷、甲醇、一氧化碳、苯和甲苯,共 18000 条样本。我们以 1 Hz 的采样率对所有信号进行重新采样,因此每个输入样本包含 260 个时间戳。同时本文还对每个样本在时间维度上进行标准化,以提高网络的学习效果。

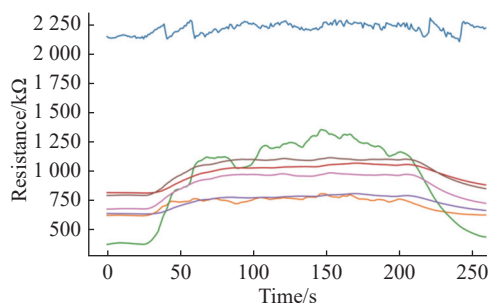


图 6 Open Sampling 数据集的代表性样本

Fig. 6 Representative sample of Open Sampling dataset

Drift 数据集是在 2007 年 1 月至 2011 年 2 月 (36 个月)期间在加州大学实验室中收集的,包含来自 16 个化学传感器的 13910 个测量值,这些传感器模拟了漂移补偿,用于对 6 种不同浓度水平的气体进行分类任务。传感器的响应以每个传感器有源层电阻的形式读出,因此每次测量都会产生一个 16 通道的时间序列。此数据集包括 6 种不同纯气态物质的记录,即氨、乙醛、丙酮、乙烯、乙醇和甲苯,每种物质的体积分数范围为 $5 \times 10^{-6} \sim 1 \times 10^{-3}$ 。原始数据被分为 10 个批次,每个批次包含如表 1 所示的每个月的测量数据。这种数据组织方式是为了确保在训练模型时,每个类和每个月都有足够的、尽可能均匀分布的数据。

Twin 数据集包括来自 8 个传感器阵列的多次重复记录。每个阵列包含 8 个 MOX 传感器,并配备定制设计的电子设备,用于控制传感器的工作温度和信号采集。所有阵列均采用相同的实验方案,测试 5 个双单元传感器的反应。每天测试一个不同的单元,每个单元随机暴露于 40 种不同的气体状态,被测装置暴露于 10 种浓度的乙醇、甲烷、乙烯和一氧化碳中。每次实验持续 600 s,每个传感器的数据以 100 Hz 的频率进行采集。

表 1 Drift 数据集的批次和月份 ID 对应表

Table 1 Mapping table of batch and month ID for Drift dataset

Batch	Month ID
1	1,2
2	3,4,8,9,10
3	11,12,13
4	14,15
5	16
6	17,18,19,20
7	21
8	22,23
9	24,30
10	36

3.2 实验设置

本文使用 Pytorch1.13.0 框架进行编程,实验在单张 NVIDIA GeForce RTX 3090 GPU 的硬件平台上进行。TCN 的通道数分别设置为 [4,8,16],膨胀因子分别设置为 [1,2,4],TCN 中的膨胀因果卷积的卷积核都设置为 (1,3),TE 模块的多头注意力的 Head 数量设置为 2,损失函数中中心损失的权重 λ 为 0.02。使用 SGD 优化器优化参数,初始学习率设置为 0.001,固定步长衰减调整学习率。3 个数据集都以 3:1:1 的比例划分训练集、测试集和验证集。根据 3.1 节中所描述的数据集规模和复杂度,以及在训练过程中对模型性能变化的观察,最终确定在 Open Sampling 数据集上训练 100 轮,在 Drift 数据集上训练 50 轮,在 Twin 数据集上训练 50 轮。

3.3 评估指标

为了评价所提算法的性能,本文使用准确率 (Accuracy)、精确率 (Precision)、召回率 (Recall)、F1-score 作为评价指标。准确率是指模型正确预测的样本数占总样本数的比例,精确率是指在所有被分类为正例的样本中真正例的比例,召回率是指在所有实际为正例的样本中被正确分类为正例的比例,F1-score 是精确率和召回率的调和平均值。计算公式分别为:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (8)$$

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

其中, TP(True Positive)表示真正例数, TN(True Negative)表示真负例数, FP(False Positive)表示假正例数, FN(False Negative)表示假负例数。

3.4 实验结果

为了验证所提方法的性能,在 3 个电子鼻数据集上使用传统机器学习方法(朴素贝叶斯(Bayes)、支持向量机(SVM)、随机森林等)、传统深度学习方法(Resnet18、ViT 等)以及气体识别领域的主流方法(CLSTM^[33]、TTCN^[34]、MCNA^[35]等)进行对比实验,结果如表 2、表 3 和表 4 所示,其中黑体表示最优结果。

在 Open Sampling 数据集上,所提方法在各项指标上都取得了最优结果,准确率达到 99.47%。机器学习方法中,朴素贝叶斯和 KNN 取得 49.95% 和 56.04% 的准确率, SVM、Logistic 回归、DT 分别取得 86.58%、85.63%、80.22% 的准确率, XgBoost 和 RF 分别取得 96.62% 和 94.42% 的准确率。传统深度学习方法中, VGG16 和 Densenet 效果最好,分别达到 96.95%、96.80% 的准确率。在气体识别任务的主流方法中, MCNA 和 Olfactory-Resnet 分别达到了 97.81% 和 98.80% 的准确率。

Drift 数据集分为 Batch 1~Batch 10 共 10 个子集,每个子集中的样本数量有 160~3 600 条不等,因此以模型在 10 个子集上的平均准确率作为在 Drift 数据集上的最终结果。机器学习方法中,效果最佳的是 RF,达到 98.60% 的准确率。深度学习方法中, Resnet 和 CAtt-CNN 取得了最好效果,分别达到了 99.43% 和 99.51% 的准确率。在 Drift 数据集上,所提出的 MTCNA 尽管在 Batch1、Batch3、Batch8 上没有达到最优,但与最优方法的结果十分接近,且在最终的平均准确率指标上仍取得了最优结果,领先于次优的 CAtt-CNN。

在 Twin 数据集上,将所提方法与传统深度学习方法 and 气体识别中的主流方法进行了对比,所提方法取得最优性能,准确率达到 99.22%。传统深度学习方法中, Resnet18 和 Mobilenet 效果最好,分别达到 94.53%、93.75% 的准确率。在气体识别任务的主流方法中, CLSTM 和 MCNA 分别达到了 98.44% 和 96.09% 的准确率。

实验结果表明,在以上 3 个数据集上本文提出的 MTCNA 都达到了最优性能,这验证了本文所提方法的识别能力,且有较强的鲁棒性。

3.5 实验分析

3.5.1 特征可视化 t-SNE 是一种非线性降维机器学习

表 2 Open Sampling 数据集上分类结果

Table 2 Classification results on Open Sampling dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
Bayes	49.95	52.57	49.53	51.00
SVM	86.58	86.23	86.15	86.19
XgBoost	96.62	96.59	96.53	96.56
Logistic regression	85.63	85.58	85.38	85.48
KNN	56.04	56.60	55.61	56.10
DT	80.22	79.88	79.84	79.86
RF	94.42	94.46	94.23	94.34
Resnet18	91.09	90.98	90.9	90.94
ViT	85.44	85.48	85.28	85.38
VGG16	96.95	97.05	96.72	96.98
Mobilenet	93.47	93.67	93.34	93.51
Densenet	96.80	96.89	96.75	96.82
CLSTM	97.68	97.68	97.65	97.67
TTCN	96.18	96.22	96.15	96.19
MCNA	97.81	97.79	97.79	97.79
ODCNN ^[36]	95.99	—	—	95.85
DWTLSTM ^[36]	90.18	—	—	89.94
Sniff-ConvNet ^[36]	95.72	—	—	95.32
Olfactory-Resnet ^[36]	98.80	—	—	98.83
Our(MTCNA)	99.47	99.46	99.47	99.47

习算法,该方法基于嵌入模型,将高维空间的数据映射到低维空间,并保留了数据的局部特征。本文使用 t-SNE 对所提取的特征进行可视化处理,图中的每一个点表示一个样本,每一种颜色表示一种类别,结果如图 7 所示。图 7(a)为原始特征图,图 7(b)所示为 MTCNA 所提取特征的可视化图。可以看到,原始特征的分布是杂乱无序的,难以直接进行分类,而提取后的特征中,相同类别的样本聚类效果显著,不同类别之间有明显的分界线。这表明时态卷积网络与 Transformer 的结合可以自动提取样本中气体类别的有效特征,有助于提高整体分类精度。

3.5.2 超参数实验 通过调整 MTCNA 的损失函数中心损失的权重 λ 、多尺度数量 k 、时态卷积网络中 TCN 块的数量 n 均能提高气体的识别能力。在本文中研究了这几个框架参数的影响。使用不同的框架参数来训练, Open Sampling 数据集上的结果如图 8 所示。

如图 8(a) 中所示,对于中心损失的权重 λ ,在 $\lambda=0.02$ 时模型取得最优性能,而随着权重的增加,模

表3 Drift数据集上的分类结果
Table 3 Classification results on Drift dataset

Method	Accuracy/%										Average
	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	
Bayes	75.6	77.2	85.8	94.1	100.0	79.2	69.1	83.3	98.9	59.5	82.27
SVM	65.6	87.6	98.1	76.5	90.2	95.2	88.8	90.0	96.9	80.9	86.98
XgBoost	96.6	98.8	99.4	97.1	97.5	99.4	99.9	96.6	97.9	99.5	98.27
Logistic Regression	96.6	98.0	99.7	100.0	97.5	99.8	99.4	96.6	98.9	99.0	98.55
KNN	95.6	98.8	99.5	97.2	97.8	98.8	99.2	95.1	99.1	96.1	97.72
DT	94.6	98.3	99.1	99.2	98.4	98.9	98.6	97.0	99.8	97.3	98.12
RF	95.6	99.3	99.7	96.4	98.8	99.0	99.9	98.0	100.0	99.3	98.60
Mobilenet	98.9	99.6	100.0	100.0	100.0	100.0	100.0	96.5	98.9	99.3	99.32
Resnet	98.9	99.6	100.0	100.0	100.0	100.0	100.0	96.5	100.0	99.3	99.43
Densenet	95.4	99.2	100.0	100.0	100.0	99.6	100.0	96.5	97.8	99.6	98.81
ViT	96.6	99.2	99.7	100.0	100.0	100.0	100.0	96.5	98.9	99.7	99.06
VGG16	98.8	99.6	100.0	100.0	100.0	100.0	100.0	96.5	100.0	99.6	99.45
TTCN	94.3	99.2	99.7	100.0	100.0	100.0	99.9	91.2	98.9	99.6	98.28
MCNA	95.6	99.2	99.4	100.0	100.0	98.9	100.0	98.3	100.0	97.6	98.90
CAtt-CNN	97.8	98.8	99.7	100.0	100.0	99.6	100.0	100.0	100.0	99.2	99.51
Our(MTCNA)	97.8	99.6	99.7	100.0	100.0	100.0	100.0	99.3	100.0	99.7	99.61

表4 Twin数据集上的分类结果

Table 4 Classification results on Twin dataset

Method	Accuracy/%	Precision/%	Recall/%	F1-score/%
Resnet18	94.53	95.28	94.53	94.8
ViT	88.28	88.65	88.28	88.47
VGG16	92.19	92.75	92.19	92.47
Mobilenet	93.75	94.33	93.75	94.04
Densenet	92.97	93.24	92.97	93.11
TTCN	95.31	95.58	95.31	95.45
MCNA	96.09	96.32	96.09	96.21
CLSTM	98.44	98.53	98.44	98.48
Our(MTCNA)	99.22	99.24	99.22	99.23

型性能逐渐下降。这说明当权重在 0.02 时, 交叉熵损失与中心损失达到平衡; 当权重系数过低时, 模型过于依赖交叉熵损失, 而忽视了同一类样本的特征之间的距离, 导致性能下降; 当权重系数过高时, 模型更加关注样本的内部结构, 而这使得模型对于训练数据的噪声或特定样本过度敏感, 从而降低了模型的泛化性能。

如图 8(b) 中所示, 当多尺度数量 k 逐渐增大时, MTCNA 的性能先升高后下降, 在 $k=3$ 时达到最佳性能。随着多尺度数量增加, 数据中的低频特征变得更加显著, 高频特征的比例降低。而在气体识别任务中, 低频信号体现的是长期的响应变化, 对模型理

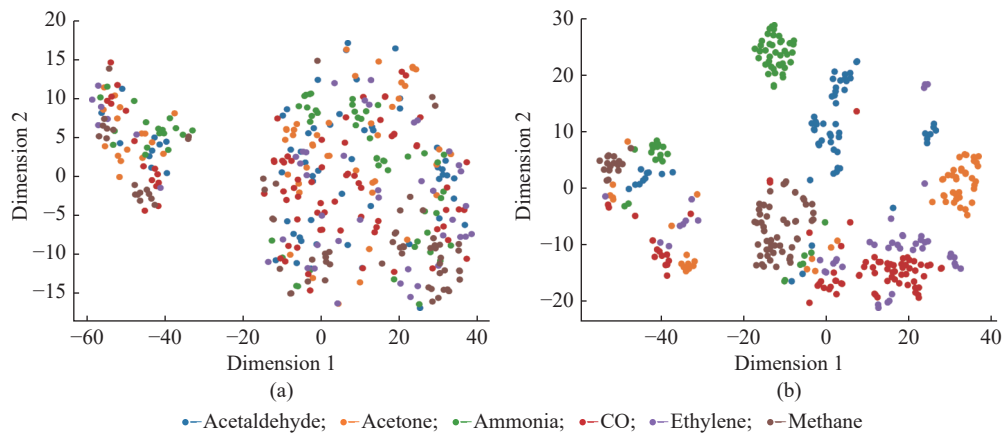


图7 原始特征可视化(a)和模型提取特征可视化(b)

Fig. 7 Original feature visualization (a) and model extraction feature visualization (b)

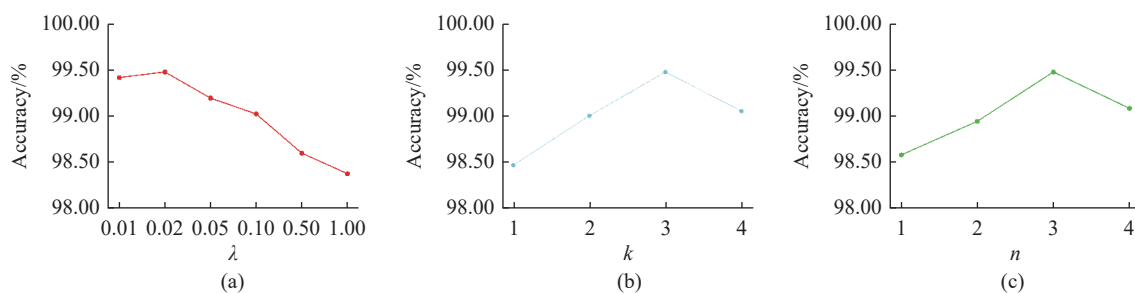


图 8 网络中不同参数下模型的效果对比图

Fig. 8 Comparison of the effects of models with different parameters in the network

解信号与气体的联系更加重要,因此模型识别精度提高。而当数据中的低频特征完全体现后,卷积网络有足够的特征来表示信号和气体之间的关系后,精度保持稳定。

如图 8(c) 中所示,当 TCN 块的数量 n 增加时,模型的性能会先升高,在 $n=3$ 时达到最高,然后逐渐下降。随着 TCN 块数量的增加,模型可以提取更多和更复杂的特征。这使得模型在初期阶段能够更好地捕获数据中的模式和结构,从而提高性能。随着 TCN 块数量的进一步增加,模型可能变得过于复杂,从而导致过拟合。模型对于训练数据中的噪声或不相关的特征过度拟合,而导致在测试集上的泛化性能下降。

3.5.3 消融实验 为了验证 MTCNA 中各模块的作用,在 Open Sampling 数据集上使用去除各模块的 MTCNA 进行了消融实验,实验结果见表 5。未添加 TCN 模块的 MTCNA 在 Open Sampling 数据集上的平均准确率下降 0.46%,未添加 Center loss 的 MTCNA 在 Open Sampling 数据集上平均准确率下降了 1.25%,而 TCN 模块和 Center loss 都未使用的 MTCNA 准确率下降了 1.66%。

表 5 MTCNA 的消融实验

Table 5 Ablation experiments of MTCNA

TCN	Center loss	Accuracy/%	Precision/%	Recall/%	F1-score/%
×	×	97.81	97.79	97.79	97.79
√	×	98.22	98.21	98.21	98.21
×	√	99.01	98.99	99.00	99.00
√	√	99.47	99.46	99.47	99.47

实验结果表明,使用时态卷积神经网络比使用传统卷积神经网络能更好地捕捉序列中的时序模式和动态变化,使模型更适用于时间序列数据。在损失函数中使用中心损失作为辅助项,能让模型在学习过程中对特征进行更好地聚类,更关注样本类间

距离,从而增强模型判别能力。通过消融实验,验证了所提方法的有效性,这些结果表明,引入时态卷积网络和中心损失对于气体分类任务的性能提升具有积极的影响。

4 结 论

本文提出了一种基于膨胀因果卷积和注意力机制的气体识别方法,旨在解决 CNN 感受野有限和不能有效提取全局特征的问题。为了应对这些挑战,设计了融合膨胀因果卷积的 TCN 模块,通过扩大网络的感受野,使模型更好地捕捉数据中的长期依赖关系。同时,使用 TE 模块提取数据中的全局特征,帮助模型关注传感器之间的潜在联系。在 3 个公开电子鼻数据集上对 MTCNA 模型进行评估,与机器学习方法和现有的气体识别领域方法进行了对比,均取得了最优性能,验证了本文方法的有效性和适用性。

参考文献:

- [1] LIN H, CHEN H, YIN C, *et al.* Lightweight residual convolutional neural network for soybean classification combined with electronic nose[J]. *IEEE Sensors Journal*, 2022, 22(12): 11463-11473.
- [2] SHI Y, YUAN H, ZHANG Q, *et al.* Lightweight interleaved residual dense network for gas identification of industrial polypropylene coupled with an electronic nose[J]. *Measurement Science and Technology*, 2023, 34(8): 085109.
- [3] 谷小婧, 林昊琪, 丁德武, 等. 基于红外气体成像及实例分割的气体泄漏检测方法 [J]. *华东理工大学学报(自然科学版)*, 2023, 49(1): 76-86.
- [4] WOJNOWSKI W, DYMERSEKI T, GEBICKI J, *et al.* Electronic noses in medical diagnostics[J]. *Current Medicinal Chemistry*, 2019, 26(1): 197-215.
- [5] MACRO S, GUTIERREZ-GALVEZ A. Signal and data processing for machine olfaction and chemical sensing: A

- review[J]. *IEEE Sensors Journal*, 2012, 12(11): 3189-3214.
- [6] DMITRZAK M, JASINSKI P, JASINSKI G. Limited selectivity of amperometric gas sensors operating in multicomponent gas mixtures and methods of selectivity improvement[J]. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 2020, 68(6): 1275-1282.
- [7] YANG J, SUN Z, CHEN Y. Fault detection using the clustering-KNN rule for gas sensor arrays[J]. *Sensors*, 2016, 16(12): 2069.
- [8] NAMAZKHAN M, ALBERS C, STEG L. A decision tree method for explaining household gas consumption: The role of building characteristics, socio-demographic variables, psychological factors and household behaviour[J]. *Renewable and Sustainable Energy Reviews*, 2020, 119: 109542.
- [9] KELLER C A, EVANS M J. Application of random forest regression to the calculation of gas-phase chemistry within the GEOS-Chem chemistry model v10[J]. *Geoscientific Model Development*, 2019, 12(3): 1209-1225.
- [10] YUAN Z, HAN E, MENG F, *et al.* Detection and identification of volatile organic compounds based on temperature-modulated ZnO sensors[J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, 69(7): 4533-4544.
- [11] ZHANG W, WANG L, CHEN J, *et al.* A novel gas recognition and concentration detection algorithm for artificial olfaction[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 9409770.
- [12] PTUCHA R, SUCH F P, PILLAI S, *et al.* Intelligent character recognition using fully convolutional neural networks[J]. *Pattern Recognition*, 2019, 88: 604-613.
- [13] LI H. Deep learning for natural language processing: Advantages and challenges[J]. *National Science Review*, 2018, 5(1): 24-26.
- [14] GU R, WANG G, SONG T, *et al.* CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2020, 40(2): 699-711.
- [15] 肖飞扬, 顾幸生. 基于并行 LSTM-CNN 的化工过程故障检测 [J]. *华东理工大学学报 (自然科学版)*, 2023, 49(3): 382-390.
- [16] WANG T, ZHANG H, WU Y, *et al.* Target discrimination, concentration prediction, and status judgment of electronic nose system based on large-scale measurement and multi-task deep learning[J]. *Sensors and Actuators B: Chemical*, 2022, 351: 130915.
- [17] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. (2018-03-01) [2018-06-01]. <https://arxiv.org/abs/1803.01271>, 2018.
- [18] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 30: 5998-6008.
- [19] MENG L, LI H, CHEN B C, *et al.* Adavit: Adaptive vision transformers for efficient image recognition[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, LA, USA: IEEE, 2022: 12309-12318.
- [20] YANG F, YANG H, FU J, *et al.* Learning texture transformer network for image super-resolution[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2020: 5791-5800.
- [21] WEN Y, ZHANG K, LI Z, *et al.* A comprehensive study on center loss for deep face recognition[J]. *International Journal of Computer Vision*, 2019, 127: 668-683.
- [22] FENG S, FARHA F, LI Q, *et al.* Review on smart gas sensing technology[J]. *Sensors*, 2019, 19(17): 3760.
- [23] YANG X, LI M, JI X, *et al.* Recognition algorithms in E-nose: A Review[J]. *IEEE Sensors Journal*, 2023, 20460-20472.
- [24] GUO J, CHENG Y, LUO D, *et al.* ODRP: A deep learning framework for odor descriptor rating prediction using electronic nose[J]. *IEEE Sensors Journal*, 2021, 21(13): 15012-15021.
- [25] XIONG Y, CHEN Y, CHEN C, *et al.* An odor recognition algorithm of electronic noses based on convolutional spiking neural network for spoiled food identification[J]. *Journal of the Electrochemical Society*, 2021, 168(7): 077519.
- [26] MA D, GAO J, ZHANG Z, *et al.* Gas recognition method based on the deep learning model of sensor array response map[J]. *Sensors and Actuators B: Chemical*, 2021, 330: 129349.
- [27] LEE J M, CHOI E J, CHUNG J H, *et al.* A DNA-derived phage nose using machine learning and artificial neural processing for diagnosing lung cancer[J]. *Biosensors and Bioelectronics*, 2021, 194: 113567.
- [28] SRINIVAS A, LIN T Y, PARMAR N, *et al.* Bottleneck transformers for visual recognition[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2021: 16519-16529.
- [29] CORDONNIER J B, LOUKAS A, JAGGI M. On the relationship between self-attention and convolutional layers[C]//*International Conference on Learning Representations*. Red Hook, NY, USA: Curran Associates, Inc, 2020.
- [30] YUAN K, GUO S, LIU Z, *et al.* Incorporating convolution designs into visual transformers[C]//*Proceedings of the IEEE/CVF International Conference on Computer Vision*. Piscataway, NJ: IEEE, 2021: 579-588.
- [31] XIAO T, SINGH M, MINTUN E, *et al.* Early convolutions help transformers see better[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 30392-30400.

- [32] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. (2018-03-19) [2021-08-25]. <https://arXiv preprint arXiv: 1803.01271, 2018>.
- [33] ZHOU K, LIU Y. Early-stage gas identification using convolutional long short-term neural network with sensor array time series data[J]. *Sensors*, 2021, 21(14): 4826.
- [34] YANG G, SONG R, WU Y, *et al*. An algorithm based on transformer and temporal convolution network for early identification of ternary gas mixture[J]. *IEEE Sensors Journal*, 2023, 23: 23753-23764.
- [35] PAN J, YANG A, WANG D, *et al*. Lightweight neural network for gas identification based on semiconductor sensor[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 71: 1-8.
- [36] ZHANG C, WANG W, PAN Y, *et al*. Curriculum learning-based approaches for end-to-end gas recognition[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 1-11

Gas Recognition Method Based on Dilatation Causal Convolution and Attention Mechanism

YU Lingwei, YANG Mengping, YANG Hai, WANG Zhe

(School of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Gas identification is of great significance in the fields of environmental monitoring, industrial safety and medical health, which can effectively detect harmful gas leaks, monitor air quality and identify disease odor markers. However, the field of gas identification is faced with the problem that the sensor data needs to be processed manually before it can be used for subsequent analysis. Convolutional neural network (CNN) has been gradually applied in gas recognition scenarios of electronic nose systems with their ability of automatic feature learning and end-to-end modeling. Although CNN performs well in this field, there are still challenges such as limited receptive field and insufficient global feature extraction, resulting in limited recognition performance. To solve these problems, a gas recognition method based on expansive causal convolution and attention mechanism is proposed. The algorithm combines the attention mechanism and multi-scale temporal convolution network in Transformer to extract global and local features, extract more representational features and obtain a larger receptive field, and capture the instantaneous information and change trend of gas. Experiments are conducted on three different data sets—Open Sampling, Drift and Twin. The results show that the proposed method achieves accuracy of 99.47%, 99.61% and 99.22%, respectively, which are superior to the existing mainstream methods, thereby confirming its effectiveness.

Key words: electronic nose system; gas identification; attention mechanism; temporal convolutional networks; time series

(责任编辑: 王晓丽)