

基于Transformer的放射性 气溶胶子体扣除算法

刘俊伟¹, 刘 涛^{2*}, 祝远帅², 吕 胜², 周宗慕^{2,3}

¹阳江核电有限公司, 广东 阳江

²中广核久源(成都)科技有限公司, 四川 成都

³西华大学电气与电子信息学院, 四川 成都

收稿日期: 2024年11月1日; 录用日期: 2024年11月11日; 发布日期: 2025年1月21日

摘 要

在人工放射性气溶胶浓度监测过程中, 氡及其子体的信号干扰是一个严重的问题。传统方法在适应环境本底变化方面表现出明显的局限性, 这些方法往往难以准确捕捉因环境变化引起的信号时变特性及其在不同时间尺度上的特征变化。本文引入了Transformer架构到气溶胶人工放射性浓度计算方法中, 利用其多头自注意力机制和特征融合模块来优化信号处理, 有效区分氡子体信号及其变化。同时, 模型通过多尺度注意力机制增强了对信号动态特性的适应性和鲁棒性。此外, Transformer架构中的特征融合模块动态调整特征表示可根据输入信号的实时变化, 进一步提升监测的准确性。实验结果显示, 引入Transformer架构后, 模型的整体性能得到显著提升, 尤其在复杂和动态的环境中处理信号效果突出, 使用该方法不仅提高了检测的准确性, 而且有效地降低了误报和漏报率。在区分氡子体信号和人工放射性信号的任务中, Matthews相关系数(MCC)达到了0.85, 整体准确率为92%。

关键词

气溶胶, 氡子体, Transformer

Transformer-Based Algorithm for Radioactive Aerosol Sub-Body Deduction

Junwei Liu¹, Tao Liu^{2*}, Yuanshuai Zhu², Sheng Lyu², Mushuai Zhou^{2,3}

¹Yangjiang Nuclear Power Co. Ltd., Yangjiang Guangdong

²CGN Jiuyuan (Chengdu) Science and Technology Co. Ltd., Chengdu Sichuan

³College of Electrical and Electronic Information, Xihua University, Chengdu Sichuan

Received: Nov. 1st, 2024; accepted: Nov. 11th, 2024; published: Jan. 21st, 2025

*通讯作者。

文章引用: 刘俊伟, 刘涛, 祝远帅, 吕胜, 周宗慕. 基于 Transformer 的放射性气溶胶子体扣除算法[J]. 核科学与技术, 2025, 13(1): 80-89. DOI: 10.12677/nst.2025.131009

Abstract

In the monitoring process of artificial radioactive aerosol concentration, signal interference from radon and its progeny poses a significant challenge. Traditional methods exhibit clear limitations in adapting to changes in environmental background, often struggling to accurately capture the time-varying characteristics of signals caused by environmental changes and their feature variations across different time scales. This paper introduces the Transformer architecture into the calculation method for aerosol artificial radioactivity concentration, utilizing its multi-head self-attention mechanism and feature fusion module to optimize signal processing and effectively differentiate radon progeny signals and their variations. At the same time, the model enhances adaptability and robustness to the dynamic characteristics of signals through a multi-scale attention mechanism. Additionally, the feature fusion module in the Transformer architecture can dynamically adjust feature representations based on real-time changes in the input signal, further improving monitoring accuracy. Experimental results show that after introducing the Transformer architecture, the overall performance of the model is significantly improved, especially in processing signals in complex and dynamic environments. This method not only increases detection accuracy but also effectively reduces false positives and missed detections. In the task of distinguishing radon progeny signals from artificial radioactive signals, the Matthews correlation coefficient (MCC) reached 0.85, with an overall accuracy of 92%.

Keywords

Aerosol, Radon Progeny, Transformer

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在核电站、核燃料和核医学等领域中,放射性气溶胶的存在是不可避免的。尤其是在核电站的日常运行和意外事故监测中,放射性物质可能以气溶胶形式释放,威胁周围环境和人群的健康[1] [2]。因此,放射性气溶胶的监测是确保辐射安全的关键环节,对减少放射性危害具有重要意义。

在气溶胶监测中,环境中的氡及其子体(如钋-218、钋-214 和铋-212)通过其放射性衰变发出的 α 和 β 射线[1],会干扰人工放射性气溶胶的测量,导致测量结果不准确,扣除氡及其子体的影响在气溶胶测量中是必不可少的。目前氡钍子体扣除方法主要有 α/β 比值法、伪 β 符合法、多能区比例系数扣除法和能谱拟合剥离法[3]。固定能区比例系数扣除法是目前应用最为广泛的算法[4]。但它依赖于预先设定的比例系数,这些系数通常基于实验室条件或历史数据确定,与实际测量条件存在差异,导致结果的可重复性和一致性差,影响结果的准确性并且提高了误报率。

21 世纪初期,随着机器学习算法在众多领域的研究和应用,它在环境氡测量和气溶胶监测中的应用也崭露头角[5]。2010 年刘仲达,蔡欣欣等人通过对厦门市东孚水氡的观测,提出采用多层神经网络(LNN)来消除环境参数(降雨量、湿度、气压、温度)对氡浓度的变化[6]。2017 年陈立等提出利用聚类分析对谱线进行分类,利用神经网络对每个类进行计算并扣除子体干扰,极大地提高了子体扣除的准确性,进而降低了人工放射性气溶胶现场监测仪的检测限[4]。2019 年龚朋彬等人提出了用最大熵原则代替传统 BP 神经网络均方差准则,建立了最大熵 BP 神经网络,创新性地将其应用对环境氡浓度预测研究中。通过仿

真实验证明改进模型平均预测误差为 5.22%，具有较强泛化能力，可以很好地用于环境氡浓度预测研究[5]。2022 年 Oni 和 Aremu 等开发了人工神经网络(ANN)模型，用于预测气象和地质参数影响下的室内氡浓度，证明了人工神经网络有助于立即评估氡超标的相关风险水平[7]。2023 年 Al-Shboul 等人探讨了人工神经网络(ANNs)和基因表达编程(GEP)模型在土壤氡呼出速率中的应用潜力，通过模型训练结果对比证明了人工神经网络在计算土壤氡呼出速率方面的优势[8]。同年，阮灵盼和陈春花等人提出了一种基于堆叠式 LSTM 网络的核运输爆炸事故放射性核素浓度预测方法。可应对突发核事故场景中放射性气溶胶扩散数据，训练最终模型预测稳定性可以达到平均绝对百分比误差(MAPE)低于 5%的 ^{239}Pu 核素浓度预测效果[9]。

综上所述，机器学习及人工神经网络可以很好地用于分析和判断气溶胶监测中氡及其子体的干扰。传统的扣除方法通常因为缺乏对环境动态变化的适应能力而导致监测精度不足。如固定能区比例系数扣除法，往往依赖于预设的比例系数，这些系数基于理想化或历史条件下的实验数据，而忽略了现场实际测量条件的多变性[4]。特别是在处理来自氡及其子体的信号时，它们无法有效区分氡子体信号与其他放射性背景噪声，进而影响了数据的可靠性和决策的准确性。针对传统方法展现出的缺陷，基于机器学习的算法能够很好地进行弥补，并且已经在放射性气溶胶监测以及氡浓度测量等方面展现出积极有效的表现，促进了放射性核素监测的发展和进步。

Transformer 是一种不同于 RNN 或 CNN 等传统模型神经网络架构的新模型，该架构因其出色的长程依赖处理能力和对时序数据的敏感性而被认为是解决此类问题的理想选择[10]。该模型通过多个不同的视角分析能谱数据，能够捕捉到氡子体信号的微妙变化，识别出这些信号与其他背景噪声的差异，从而达到精准排除的效果。本文通过在多个实际场景中测试，验证了该 Transformer 模型能够显著提高气溶胶监测的准确性和效率。实验结果证明，该模型在区分氡子体信号与其他放射性背景时，表现出优越的性能，特别是在动态变化的环境中，能够实现对复杂信号快速而准确地解析

2. 材料与方法

2.1. 固定能区比例系数扣除法

固定能区比例系数扣除法主要分为两种方法：二能区法和多能区法。相较于二能区法，多能区法能够更细致地考虑不同能量水平下氡子体 α 粒子的干扰程度，从而实现更为精确的干扰扣除[4]。目前，多能区法已成为主流技术。在多能区法中，能区的划分通常包括人工能区和氡子体能区。进一步地，根据不同氡子体对测量结果的贡献，氡子体能区可被细分为多个子能区。值得注意的是，不同的氡子体对人工放射性核素测量的干扰系数存在差异。

比例系数扣除法的关键在于准确确定子体能区对人工能区的干扰计数及其相应的系数。通过在人工智能区中精确扣除由子体引起的干扰，可以确保人工核素的检测下限得到有效保障。不同能区的划分范围及其来源详见图 1 所示。

扣除方法描述：某氡子体在人工核素能区产生的拖尾计数与其峰计数的比值。 ^{214}Po 有两个拖尾比例系数 $k_{1,3}$ 和 $k_{2,3}$ 分别表示其在一和二能区产生的拖尾计数与其在三能区产生的峰计数的比值。 ^{218}Po 有一个拖尾比例系数 $k_{1,2}$ 表示其在一能区产生的拖尾计数与其在二能区产生的峰计数比值。那么人工 α 核素产生的净计数 N_0 可以表示为：

$$N_0 = N_1 - k_{1,3}N_3 - k_{1,2}(N_2 - k_{2,3}N_3) \quad (1)$$

其中， N_1 、 N_2 和 N_3 分别表示一、二、三能区的总计数；其中 $N_2 - k_{2,3}N_3$ 表示 ^{218}Po 的峰计数。

通过确定各能区扣除比例系数值 $k_{m,n}$ ，就可以计算氡、钍及其子体在人工放射性核素能区中所占

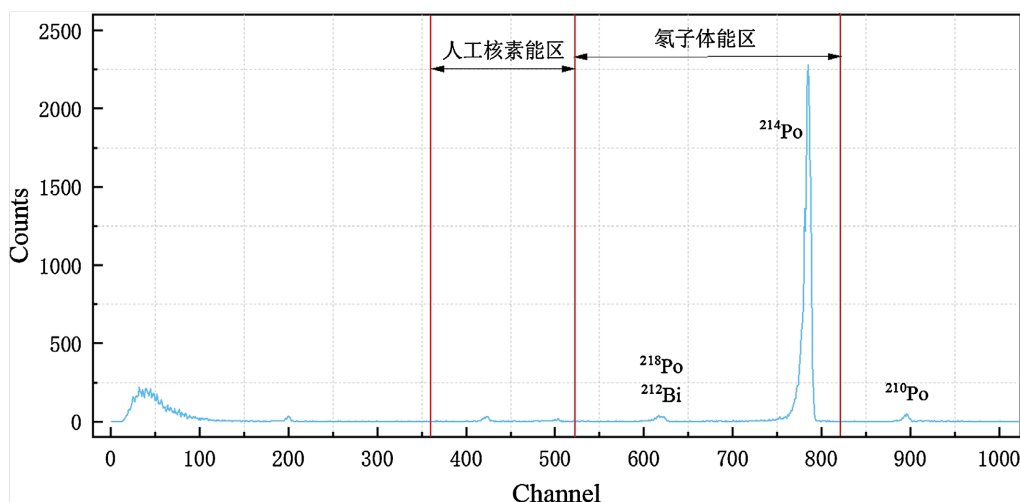


Figure 1. Energy region division diagram

图 1. 能区划分图

比例。测得人工放射性核素能区计数通过扣除干扰核素的相应计数就可以得到准确的人工放射性核素监测数据。

2.2. Transformer 模型

Transformer 技术代表了一种革命性的人工智能模型，其核心思想是构建一种深度神经网络模型，仿效人类语言理解与生成的方式，实现对序列数据的高效处理和建模[11]。与传统的循环神经网络相比，Transformer 引入了注意力机制，从而在处理长距离依赖关系时表现更为出色。相较于传统的循环神经网络，Transformer 模型采用了自注意力机制来捕捉输入序列中各个位置之间的关系，这使得模型能够并行处理序列数据，从而大幅提高了训练速度和效率。Transformer 模型的设计使得其在机器翻译、语言建模、文本生成等各种任务中占据领先地位。此外，Transformer 还包括用于将输入序列编码成隐藏表示的编码器和将隐藏表示解码成输出序列的解码器，这种结构使得模型在处理翻译等任务时表现出色。

2.2.1. 数据预处理与特征提取

在输入数据处理前，需要对数据进行预处理和特征提取，包括小波变换滤波和背景辐射信号的获取。将数据进行标准化处理，具体使用 Z-score 标准化方法。将数据转换为均值为 0，标准差为 1 的标准正态分布。然后应用小波变换对数据进行滤波处理，选用 Daubechies 小波基函数(db4)，分解级数为 5 级，以去除潜在的高频干扰并增强信号特征。

小波变换滤波是为了去除高频噪声和低频背景噪声，采用小波变换进行滤波。具体步骤如下：

(1) 小波分解：将原始信号分解为不同尺度的频率成分。

$$\{cA_n, cD_1, cD_2, \dots, cD_n\} = \text{DWT}(X) \quad (2)$$

其中， cA_n 是近似系数， cD_i 是第 i 级的细节系数。

(2) 阈值处理：对高频细节系数进行软阈值处理以去除噪声。

$$cD_i = \begin{cases} \text{sign}(cD_i)(|cD_i| - \lambda) & \text{if } |cD_i| > \lambda \\ 0 & \text{if } |cD_i| \leq \lambda \end{cases} \quad (3)$$

其中， λ 为阈值。

(3) 小波重构：将处理后的系数组合，重构滤波后的信号 X_{denoised}

$$X_{\text{denoised}} = \text{IDWT}(cA_n, cD_1, cD_2, \dots, cD_n) \quad (4)$$

经过滤波处理后，需要从数据中提取背景辐射信号。背景辐射信号的获取方法如下：

(1) 统计特征提取：计算滤波后信号的统计特征

$$F_{\text{stat}} = [\text{mean}(X_{\text{denoised}}), \text{std}(X_{\text{denoised}}), \max(X_{\text{denoised}}), \min(X_{\text{denoised}})] \quad (5)$$

(2) 峰值检测：使用峰值检测算法识别信号中的主要峰值及其特征

$$F_{\text{peaks}} = \text{findpeaks}(X_{\text{denoised}}) \quad (6)$$

(3) 低能拖尾效应：计算低能区域(前 100 个通道)的总计数。

$$\text{LET} = \sum_{i=1}^{100} X_{\text{denoised}}[i] \quad (7)$$

最终的特征向量为：

$$F = [F_{\text{stat}}, F_{\text{peaks}}, \text{LET}] \quad (8)$$

2.2.2. 输入嵌入

经过预处理和特征提取后，特征向量通过线性嵌入层转换为高维特征空间：

$$E = FW_{\text{emb}} + b_{\text{emb}} \quad (9)$$

其中， $W_{\text{emb}} \in \mathbb{R}^{d_{\text{input}} \times d_{\text{model}}}$ 和 $b_{\text{emb}} \in \mathbb{R}^{d_{\text{model}}}$ ， d_{input} 是输入特征的维度， d_{model} 是模型的维度，设为 256。

使用预训练的词嵌入模型，将输入数据转化为 256 维的高维向量；引入位置编码以保留时间序列信息，使用正弦和余弦函数生成位置编码；编码器和解码器各包含 6 层，每层由多头自注意力机制(8 个头)和前馈神经网络(2048 个隐藏单元)组成；使用 Adam 优化器进行参数优化，初始学习率设为 0.001，结合学习率预热及逐步衰减策略；引入 0.1 的 dropout 正则化以防止过拟合；采用带权重调整的交叉熵损失函数，以应对类别不平衡的问题。位置编码与输入嵌入的集成方式，强调了模型如何通过位置信息增强对数据的学习。具体结构如下图 2 所示。

2.2.3. 位置编码

为了捕捉数据的时序性，位置编码被添加到嵌入表示中：

$$H_0 = E + P \quad (10)$$

位置编码 P 计算如下：

$$\text{PE}_{(\text{pos}, 2i)} = \sin\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right) \quad (11)$$

$$\text{PE}_{(\text{pos}, 2i+1)} = \cos\left(\frac{\text{pos}}{10000^{2i/d_{\text{model}}}}\right) \quad (12)$$

2.2.4. Transformer 编码器

编码器通过多层处理位置编码后的输入，每层均包含多头自注意力机制和前馈神经网络：

$$H_{l+1} = \text{LayerNorm}(H_l + \text{MultiHead}(H_l, H_l, H_l)) \quad (13)$$

多头自注意力机制的具体实现如下：

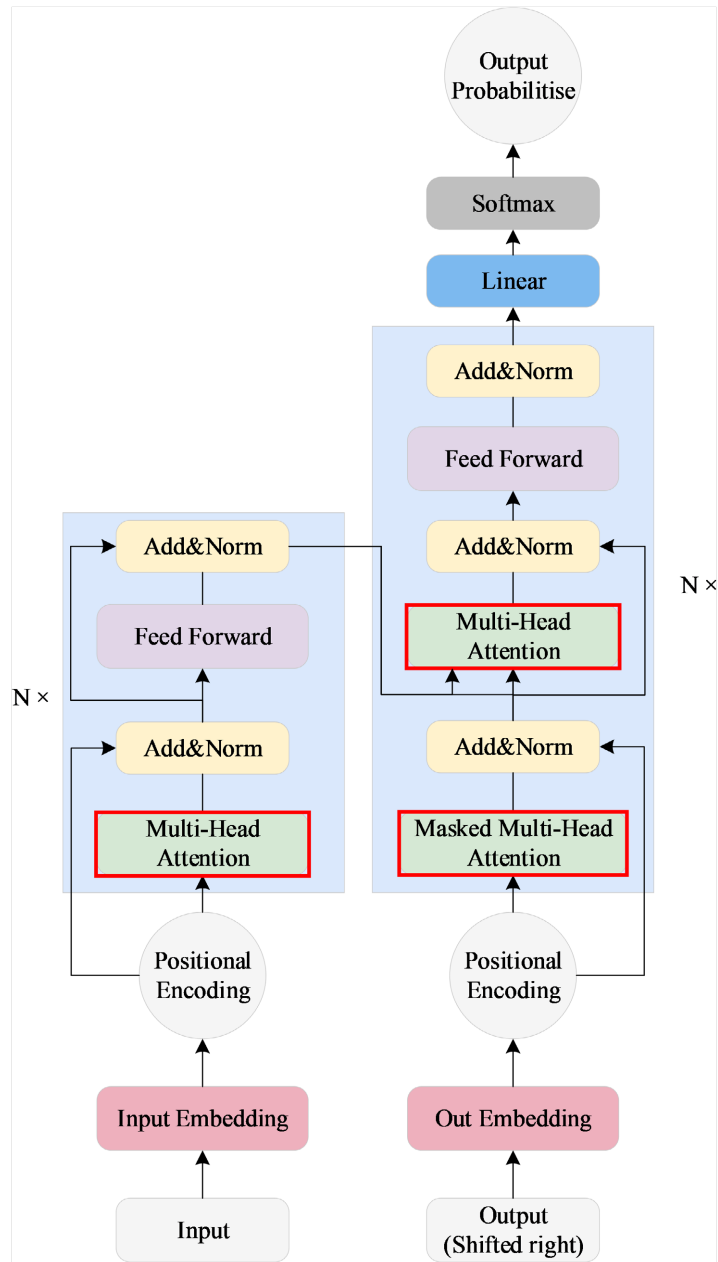


Figure 2. Internal structure diagram of Transformer

图 2. Transformer 的内部结构图

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_h) \mathbf{W}^O \quad (14)$$

$$\text{head}_i = \text{Attention}(\mathbf{Q} \mathbf{W}_i^Q, \mathbf{K} \mathbf{W}_i^K, \mathbf{V} \mathbf{W}_i^V) \quad (15)$$

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q} \mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} \quad (16)$$

为了捕捉数据中的多尺度特征，本章节设计了一个多尺度注意力机制，通过并行计算不同时间尺度上的注意力分布，有效捕捉短期和长期的信号特征，从而增强对氦子体信号的识别能力。

编码器内部结构包括的工作原理, 包括多头注意力机制如何在不同尺度上处理光谱数据, 捕捉重要的信号特征。

2.2.5. 特征融合与分类

编码后的特征通过特征融合模块进行整合, 该模块结合了深度学习特征和从原始光谱数据中提取的物理特征(如峰值位置和强度、低能拖尾效应等), 以提高模型的预测精度。特征融合模块采用以下公式:

$$\mathbf{F}_{\text{fused}} = \text{Concat}(\mathbf{H}_{\text{encoder}}, \mathbf{F}_{\text{engineered}}) \quad (17)$$

其中, $\mathbf{H}_{\text{encoder}}$ 是编码器输出的深度特征, $\mathbf{F}_{\text{engineered}}$ 是从原始数据中提取的物理特征。融合后的特征通过全连接层进行分类。

2.2.6. 背景辐射信号的提取

在采样测量人工放射性气溶胶时, 氡、钍子体放射性气溶胶会连同人工放射性气溶胶一同被采集到采样滤纸上。其中, 背景辐射信号属于天然本底带来的噪声, 会根据测量条件不同而变化, 属于需要简化并剔除的相关干扰。一旦确定监测环境, 本底辐射(非氡、钍核素)在后续的监测过程中几乎不会改变。而环境中的氡、钍子体浓度随地理条件、时间、和气象参数的变化会在较大范围内变化[12]。为了达到区分目的, 需要精确识别氡、钍子体的贡献值, 才能提高原始数据的判定准确度和检测下限。特征融合实现如下:

$$\mathbf{O} = \text{softmax}(\mathbf{F}_{\text{fused}} \mathbf{W}_{\text{out}} + \mathbf{b}_{\text{out}}) \quad (18)$$

在完成特征融合与分类后, 我们使用分类器的输出来提取背景辐射信号和氡子体信号。两种信号分类提取具体步骤如下:

- (1) 分类结果处理: 根据分类结果, 将信号分为氡子体信号和背景辐射信号。具体实现如下:

$$\mathbf{C}_{\text{radon}} = \mathbf{O}[:, 1] \odot \mathbf{X}_{\text{denoised}} \quad (19)$$

$$\mathbf{C}_{\text{background}} = \mathbf{O}[:, 0] \odot \mathbf{X}_{\text{denoised}} \quad (20)$$

其中 $\odot$ 表示按元素乘积, $\mathbf{O}[:, 1]$ 和 $\mathbf{O}[:, 0]$ 分别表示分类器输出为氡子体信号和背景辐射信号的概率。

- (2) 背景辐射信号计算: 通过对背景辐射信号部分进行积分, 得到最终的背景辐射信号 $C_{\text{background}}$:

$$C_{\text{background}} = \sum_{i=1}^{1024} \mathbf{C}_{\text{background}}[i] \quad (21)$$

Transformer 是一种不同于 RNN 或 CNN 等传统模型神经网络架构, 引用残差连接和层归一化等技术来加速并行计算。Transformer 模型中多头注意机制的并行计算大大提高了训练和推理的效率, 允许更大的模型和更长的序列处理。模型内部注意力机制和多头注意力如下图 3 所示。

3. 结果与讨论

本研究使用合成数据集进行实验, 该数据集包括从低至 50 Bq/m³ 到高达 10,000 Bq/m³ 不等的氡浓度。数据集模拟了实际环境监测中可能遇到的各种情况, 如不同氡释放速率、环境变化和灵敏度等因素的影响。实验样本共 2000 个, 被随机分为三个部分: 70% 作为训练集, 用于模型训练; 15% 作为验证集, 用于模型调优和超参数选择; 15% 作为测试集, 用于模型的最终性能评估。所有实验均在配置有 NVIDIA GeForce RTX 2060 SUPER GPU 的计算平台上完成。选择准确率、Matthews 相关系数(MCC)和最低检测限(MDL)作为主要评价指标。

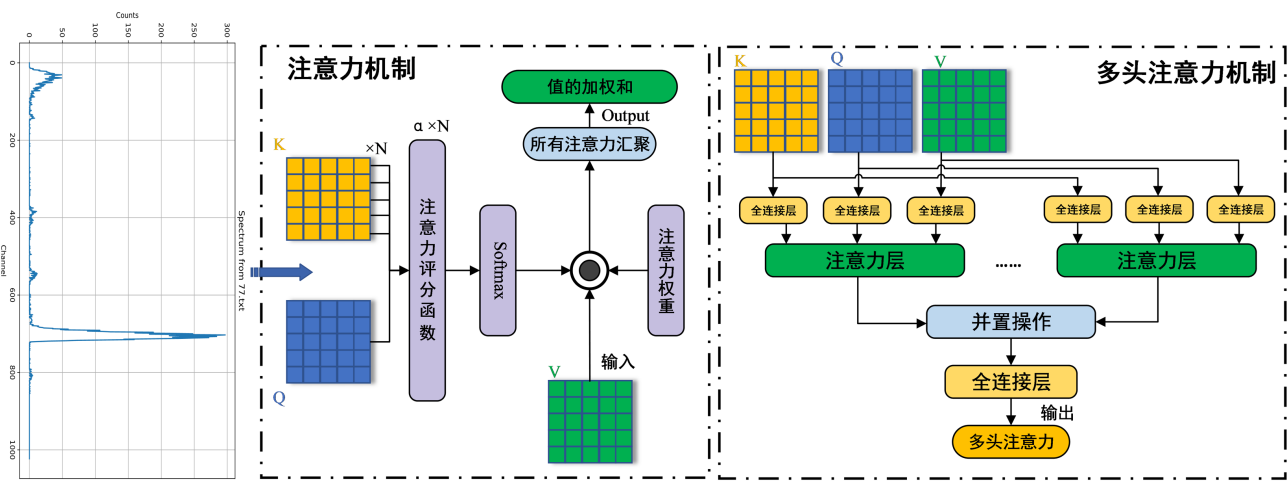


Figure 3. Attention mechanism and multi-head attention architecture diagram
图 3. 注意力机制和多头注意力结构示意图

3.1. 训练模型性能

为了对新构建的模型的性能进行评价, 本文选择了近年来在测氡领域性能优秀的三个基于机器学习的算法结构为比较对象, 它们分别是支持向量机、随机森林、卷积神经网络[4][5]。采用同样的数据样本进行训练, 各训练模型预测结果如下表 1 所示。

Table 1. Model performance comparison
表 1. 模型性能比较

模型	准确率(%)	MCC	MDL ($\times 10^{-4}$ Bq/m ³)	超参数设置
支持向量机	80.56	0.60	20	核函数: RBF, C = 1, gamma = 0.1
随机森林	84.10	0.65	15	树数: 100, 最大深度: None
卷积神经网络	88.51	0.72	10	隐藏层: 3 层, 每层 128 个神经元
本研究模型	92.23	0.85	5	见 2.2. 节详细设置

表 1 展示了本研究中的 Transform 模型与其他三类常见算法结构的比较结果。表 1 的实验结果显示, Transformer 模型在准确率(Accuracy)、Mathews 相关系数(MCC)和最小检测限(MDL)上均优于对比模型, 说明模型对于氡浓度的低阈值检测具有较高的灵敏度和准确性, 表明在放射性气溶胶子体扣除任务中的显著优势。

从准确率来看, 支持向量机和随机森林的准确率分别为 80.56%和 84.10%, 卷积神经网络的准确率提升至 88.51%, 而本研究的 Transformer 模型则达到了 92.23%。这一显著提升表明, Transformer 模型能够更有效地捕捉气溶胶数据中的复杂特征, 并在预测时更具准确性。其原因在于 Transformer 的自注意力机制能够在全局范围内学习特征关系, 从而在序列数据处理中表现出色。

MCC 是一种在类别不平衡情况下衡量分类性能的指标。实验中, 支持向量机和随机森林的 MCC 分别为 0.60 和 0.65, 卷积神经网络为 0.72, 而 Transformer 模型达到了 0.85, 表现出更强的正负样本区分能力。Transformer 的多头自注意力机制能够更细致地关注数据中不同特征的重要性, 从而在正负样本的预测上更加平衡。

本研究模型的 MDL 为 5×10^{-4} Bq/m³, 低于其他模型(如 SVM 的 20×10^{-4} Bq/m³ 和 CNN 的 10×10^{-4}

Bq/m³)。MDL 的降低意味着模型具有更高的检测灵敏度和噪声适应能力。Transformer 通过位置编码(positional encoding)能够更好地捕捉气溶胶信号的时间序列特征，增强了对低强度信号的响应能力。

在超参数设置上，本研究对比了各模型的关键超参数。支持向量机的核函数选用 RBF 核，而随机森林树数采用 100，卷积神经网络使用三层 128 神经元的隐藏层。相比之下，Transformer 模型在多头自注意力和编码器结构上进行了优化，使其在大规模数据的特征提取上具有独特优势。此外，卷积神经网络虽然在卷积操作上有助于提取局部特征，但在处理长序列时存在局限性，难以捕捉长距离的特征关系。这些实验结果与文献中对不同模型的特性描述一致，进一步证明了 Transformer 架构在序列数据分析中的适用性。

综上所述，Transformer 模型在放射性气溶胶子体扣除中的优异表现验证了其作为深度学习模型的先进性和科学性。未来的研究可以考虑在真实环境数据上进一步验证本模型，以增强其在实际应用中的可靠性。

3.2. 训练结果对比

Transformer 基础架构提供的多头自注意力机制和位置编码在处理环境监测数据时至关重要。特别是位置编码，对于捕捉时间序列数据中的局部依赖和全局预测非常有效。此外，消融实验进一步证明了这些组件的必要性，去除任何一个都会显著降低模型的准确性和最低检测限。详细结果见下表 2。

Table 2. Results of the ablation experiment
表 2. 消融实验结果

组件组合	准确率(%)	MCC	MDL ($\times 10^{-4}$ Bq/m ³)
完整模型	92	0.85	5
-小波滤波	89	0.80	8
-位置编码	88	0.79	10
-多头自注意力机制	87	0.76	12
-小波滤波和位置编码	86	0.74	14
-小波滤波和多头自注意力机制	85	0.72	15
-位置编码和多头自注意力机制	84	0.70	16

通过消融实验可以观察到，本模型各组件对性能有显著影响。去除位置编码后，模型准确率和 MCC 显著下降，最低检测限升至 0.001 Bq/m³，表明位置编码对捕捉时间序列数据中的局部依赖和全局上下文至关重要。去除多头自注意力机制，模型性能显著下降，MCC 降至 0.76，最低检测限升至 0.0012 Bq/m³，显示出其在捕捉不同位置的相关性和增强特征提取能力方面的重要性。去除小波滤波，模型性能下降，尤其在噪声环境中准确性减弱，最低检测限升至 0.0008 Bq/m³，说明小波滤波对去除背景噪声和提高信号清晰度有关键作用。在组合实验中去除位置编码和多头自注意力机制时，模型性能最差，MCC 降至 0.70，最低检测限升至 0.0016 Bq/m³。综上，位置编码、多头自注意力机制和小波滤波对提高模型性能尤为重要，尤其在处理噪声时间序列数据时。

4. 总结

本研究首次将基于 Transformer 的核素识别模型应用于放射性气溶胶子体扣除任务，展示了其在提高监测准确性和降低检测限方面的显著优势。Transformer 模型通过其独特的自注意力机制，实现了对气溶

胶数据中复杂特征的有效捕捉,特别是在处理动态变化的环境信号时表现出色。实验结果表明,该模型在准确率、MCC 和 MDL 等关键指标上均优于传统方法和其他机器学习模型,证明了其在放射性气溶胶监测领域的先进性和科学性。

尽管 Transformer 模型在本研究中取得了令人鼓舞的成果,但目前仍存在一些局限性。例如,模型的泛化能力需要在更广泛的数据集上进行验证,以确保其在不同环境条件下的稳定性。在特定的使用场景,可能需要进行参数修正以提高模型的识别准确度。此外,模型的计算复杂度较高,可能在资源受限的环境中部署时面临挑战。未来的工作将集中在模型的优化、新数据集的开发以及不同应用场景的测试上,以进一步提高模型的实用性和可靠性。

在实际应用方面,本研究提出的模型有望集成到核电站、核燃料处理和核医学等领域的监测系统中,为核设施工作人员提供一种更可靠的放射性气溶胶监测工具。这不仅能够提高监测数据的准确性,还能够降低误报和漏报率,从而增强核安全性。

此外,本研究的成果也可能对其他相关领域产生启示。例如,在环境监测和医学成像等领域,Transformer 模型的自注意力机制可能有助于提高数据处理的效率和准确性。未来的研究可以探索该模型在这些领域的应用潜力,推动相关技术的发展和革新。

总体而言,本研究不仅为放射性气溶胶监测领域提供了一种新的技术手段,也为神经网络技术在核安全领域的应用开辟了新的道路。随着技术的不断进步和应用的不断拓展,我们有理由相信,基于 Transformer 的模型将在未来的核安全监测中发挥越来越重要的作用。

参考文献

- [1] 李惠彬. 高氡环境下钍气溶胶连续监测技术研究及设备研制[D]: [博士学位论文]. 北京: 清华大学, 2013.
- [2] 黄宪果, 夏文友, 涂俊等. α 放射性气溶胶连续测量的本底扣除技术及探测灵敏度研究[J]. 辐射防护, 2015, 35(2): 93-96.
- [3] 谷铁男. 人工 α 放射性气溶胶监测中能量甄别方法研究[D]: [硕士学位论文]. 北京: 清华大学, 2012.
- [4] 陈立, 顾民, 曾国强, 等. 基于神经网络的人工放射性气溶胶中氡子体扣除算法[J]. 核技术, 2017, 40(9): 41-45.
- [5] 龚朋彬, 过惠平, 段中山. 改进 BP 神经网络法预测环境氡浓度[J]. 核电子学与探测技术, 2019, 39(3): 279-284.
- [6] 刘仲达, 蔡欣欣, 卓群等. 厦门东孚井水氡异常研究[J]. 大地测量与地球动力学, 2010, 30(A1): 81-83.
- [7] Oni, O.M., Aremu, A.A., Oladapo, O.O., Agboluaje, B.A. and Fajemiroye, J.A. (2022) Artificial Neural Network Modeling of Meteorological and Geological Influences on Indoor Radon Concentration in Selected Tertiary Institutions in Southwestern Nigeria. *Journal of Environmental Radioactivity*, **251**, Article ID: 106933. <https://doi.org/10.1016/j.jenvrad.2022.106933>
- [8] Al-Shboul, K.F., Almasabha, G., Shehadeh, A. and Alshboul, O. (2023) Exploring the Efficacy of Machine Learning Models for Predicting Soil Radon Exhalation Rates. *Stochastic Environmental Research and Risk Assessment*, **37**, 4307-4321. <https://doi.org/10.1007/s00477-023-02509-x>
- [9] 阮灵盼, 陈春花, 陈黎伟, 等. 基于长短期记忆网络的涉核运输事故后果预测[J]. 辐射研究与辐射工艺学报, 2023, 41(4): 68-80.
- [10] Guo, J., Cai, Q., An, J., Chen, P., Ma, C., Wan, J., et al. (2022) A Transformer Based Neural Network for Emotion Recognition and Visualizations of Crucial EEG Channels. *Physica A: Statistical Mechanics and its Applications*, **603**, Article ID: 127700. <https://doi.org/10.1016/j.physa.2022.127700>
- [11] Vaswani, A., et al. (2017) Attention Is All You Need. arXiv: 1706.03762.
- [12] 中国科学院原子能研究所. 辐射防护计数汇编[M]. 北京: 原子能出版社, 1978.