Advances in Environmental Protection 环境保护前沿, 2026, 16(7), 1257-1271
Published Online July 2026 in Hans. https://www.hanspub.org/journal/aep
https://doi.org/10.12677/aep.2026.167127

Hans汉斯

基于注意力机制融合的CNN-LSTM-Transformer模型水质分类研究

——以甘肃省为例

蒋智超，杨占山*

青海民族大学数学与统计学院，青海 西宁

收稿日期：2026年6月2日；录用日期：2026年7月10日；发布日期：2026年7月23日

摘 要

为了解决单一机器学习模型对非线性、有周期性的水质数据的分类效果不理想的问题，本文提出了一种用注意力机制融合的CNN-LSTM-Transformer模型。本研究选取了甘肃省2023年的十个水质观测站的时序数据，结合实际进行了数据合并。在处理异常值方面，本文先对异常值进行识别，根据不同的情况采用不同的方法处理异常值；而后进行了特征筛选，选择了跟水质类别关联性最高的4个特征。在模型构建方面，通过串联加注意力机制融合了CNN-LSTM-Transformer这三个模型。实验结果表明，与传统的单一模型相比，该模型具有更好的分类效果和更小的误差，精确率达到了96.2%。消融实验进一步表明，融合后的模型的准确率，精确率，召回率均提升了10%左右，并且在其他数据集上的验证实验也展现了该模型较强的鲁棒性。通过注意力机制融合的该模型，本身具有一定的自我动态调整能力，对不同的数据都有着较好的适应性，对现实中水质分类提供了一定的理论参考意义。

关键词

异常值处理，注意力机制，CNN-LSTM-Transformer模型，水质分类

Study on Water Quality Classification Based on Attention Mechanism-Fused CNN-LSTM-Transformer Model

—A Case Study of Gansu Province

Zhichao Jiang, Zhanshan Yang*

School of Mathematics and Statistics, Qinghai Minzu University, Xining Qinghai

*通讯作者。

文章引用: 蒋智超, 杨占山. 基于注意力机制融合的 CNN-LSTM-Transformer 模型水质分类研究[J]. 环境保护前沿, 2026, 16(7): 1257-1271. DOI: 10.12677/aep.2026.167127

Received: June 2, 2026; accepted: July 10, 2026; published: July 23, 2026

Abstract

To tackle the problem of unsatisfactory classification performance of single machine learning models in processing nonlinear and periodic water quality data, this paper proposes a CNN-LSTM-Transformer model integrated via an attention mechanism. This study selects time-series data from ten water quality monitoring stations in Gansu Province in 2023 and carries out data integration in combination with practical scenarios. In terms of outlier processing, outliers are first identified and addressed with tailored methods according to different situations; feature selection is then implemented to extract four features with the strongest correlation to water quality categories. For model construction, the CNN, LSTM and Transformer models are fused through serial connection combined with the attention mechanism. Experimental results show that the proposed model delivers better classification performance and smaller errors than traditional single models, with a precision rate reaching 96.2%. Further ablation experiments demonstrate that the accuracy, precision and recall rates of the fused model are all improved by approximately 10%. In addition, validation experiments on other datasets verify the strong robustness of the model. Benefiting from the attention mechanism-based fusion, the proposed model possesses a certain self-dynamic adjustment ability and favorable adaptability to diverse datasets, providing a valuable theoretical reference for practical water quality classification.

Keywords

Outlier Processing, Attention Mechanism, CNN-LSTM-Transformer Model, Water Quality Classification

1. 引言

水资源是生态系统稳定与人类社会可持续发展的核心保障，其中水质分类是水环境监测、污染治理与水资源优化配置的关键前提。随着工业化与城市化进程加快，农业面源污染、工业排污等问题导致水质恶化风险加剧，亟需高效的水质分类方法支撑水环境管理决策。

传统水质分类主要有单因子评价法、模糊综合评价法，灰色聚类法等。徐祖信开发了单因子水质标识指数法，该方法可以完整刻画评价指标的水质类别、水质数据、与水环境功能区类别的比较情况[1]。而后，人们在单因子评价法的基础上发现用综合污染指数法更加全面。比如罗芳等就运用改进的内梅罗污染指数法以及单因子评价法做实证分析，发现改进的内梅罗污染指数法相比单因子评价法来说更加科学合理[2]。同期，邓峰首次提出了模糊综合指数法，基于模糊数学理论，通过构建隶属度矩阵描述指标对水质等级的模糊边界，结合权重向量合成综合隶属度，最终判定最优类别[3]。虽然该方法是从大气问题中提出的，但是后来人们发现该方法也适用于水质分析评价，并且后续人们对该方法进行了更多的尝试和改进。如周林飞等将模糊评价应用于湿地水质评价，验证其在复杂生态系统中的适用性[4]。侯玉婷等针对喀斯特地区水质特点，改进传统模糊评价模型，提高评价准确性[5]。谢艾玉等提出了一种基于变异系数的动态组合赋权模型，更加符合水质数据的季节周期性[6]。后面人们又用系统的观点来看待水质

数据，越来越多的人用灰色系统聚类的方法来分析水质类别。贺北方等系统构建基于灰色聚类决策的水质评价模型，详细介绍白化函数确定、聚类权计算和决策过程，提供完整算法和应用案例[7]。敖成欢等运用模糊综合法和灰色关联法更加合理的评价了水质类别[8]。

近年来，机器学习算法如随机森林、支持向量机、LSTM 等逐渐应用于水质分类。葛新越等在 BP 神经网络的基础上加上了优化算法，使得模型平均收敛速度比传统 BP 神经网络快且不易陷入局部极小值，并保证了较好的准确率[9]。方志豪等通过调整不同属性以不同的权值，削弱了朴素贝叶斯条件独立性的假设，使分类结果更接近实际类别[10]。徐玲等通过 SMOTE 算法调整数据的不平衡性，然后运用 GA 优化算法，提出了一种融合模型，大幅提高了预测的准确率[11]。项新建等运用多种分类器和证据理论结合，并与单一分类器的效果进行对比，说明融合模型比单一模型的效果更好[12]。刘凯等运用多种机器学习方法做比较，结果展示了 LSTM 在水质分类这一领域的优越性[13]。

传统水质分类方法存在权重设定主观、未充分挖掘指标关联性等局限，难以适应大规模连续监测需求。近年来，机器学习算法逐渐应用于水质分类，但一方面由于水质数据具有特征多关联性、时序周期性、易受异常值干扰，存在缺失值等特点，另一方面，单一机器学习模型泛化能力弱、单一深度学习模型难以兼顾局部与全局特征，导致效果也不是很稳定。所以本文针对水质数据特殊性和单一模型的缺点，提出“先识别再处理”的差异化异常值处理策略，适配水质数据中真实异常与记录错误并存的特点，提升数据质量；设计注意力机制融合框架，实现 CNN 局部特征、LSTM 时序特征与 Transformer 全局特征的自适应集成，强化特征表达能力；以甘肃省跨流域站点数据为样本，验证模型的跨区域泛化能力，实现水质分类的快速精准判定，为干旱半干旱地区水质分类提供技术支撑。

2. 数据来源及方法原理

2.1. 数据来源

本研究水质时序数据来源于国家地表水水质自动监测实时发布平台公开监测数据，数据获取网址为 https://szzdjc.cnemc.cn:8070/GJZ/Business/Publish/Main.html。研究筛选甘肃省境内扶和桥、什川桥、湟水桥、青城桥、五佛寺、靖远桥、桦林、葡萄园(牛背)、折桥、地沟桥共 10 处地表水监测断面 2023 年监测序列开展分析。其中，扶和桥、什川桥、青城桥、靖远桥、五佛寺属于黄河干流；湟水桥属于湟水河流域，桦林、葡萄园(牛背)属于渭河流域，折桥、地沟桥属于大夏河流域，这些河流属于黄河一级支流。该数据为每隔 4 小时观测一次的时序数据，共计 19,460 个。每个数据都包含水质类别，水温、pH、溶解氧、电导率、浊度、高锰酸盐指数、氨氮、总磷、总氮这 10 个指标。其中水质类别分为六类，分别是I类，II类，III类，IV类，V类，劣V类，详细指标描述见表 1。

Table 1. Description of main indicators for water quality data
表 1. 水质数据主要指标说明

名称	单位	描述
水温	(℃)	反映水体热量状态
PH	(无量纲)	表征水体酸碱性
溶解氧	(mg/L)	指水中溶解的氧气含量
电导率	(μS/cm)	反映水体中离子浓度
浊度	(NTU)	表征水体浑浊程度
高锰酸盐指数	(mg/L)	反映水体中可被氧化有机物与还原性无机物含量

续表

氨氮	(mg/L)	水体中氨态氮的含量
总磷	(mg/L)	水体中各类磷的总和
总氮	(mg/L)	水体中各类氮的总和

2.2. CNN 卷积神经网络

CNN 卷积神经网络是一种专为处理网格结构数据设计的深度学习模型，核心是通过“局部感知 + 参数共享”高效提取特征，广泛用于图像、语音等领域。主要包括以下几部分：

(1) 卷积层：用卷积核即小尺寸矩阵，对输入数据做滑动相乘求和，提取局部特征，且同一卷积核参数共享，减少计算量。

(2) 池化层：对卷积层输出的特征图进行下采样，如最大值池化、平均池化法，保留关键特征的同时缩小数据维度，提升模型泛化能力。

(3) 全连接层：将池化层输出的多维特征图展平为一维向量，通过全连接计算完成分类、回归等任务。

(4) 激活函数：在卷积层和全连接层后插入，引入非线性，让模型能拟合复杂数据关系。

2.3. LSTM 长短期记忆神经网络

LSTM 是一种特殊的循环神经网络，专门设计用来解决传统 RNN 在处理长序列数据时遇到的长期依赖问题，尤其是梯度消失或梯度爆炸问题。其核心在于引入了记忆单元和门控机制。使用一个专门的“记忆细胞”，贯穿整个时间序列，主要通过三个“门”，这些门赋予了 LSTM 显式地控制信息流的能力，其中：

遗忘门：决定记忆单元中哪些旧信息需要被遗忘

$$f_t = \sigma\left(w_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{1}$$

输入门：决定当前输入中哪些信息需要被存入记忆单元

$$i_t = \sigma\left(w_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{2}$$

输出门：决定当前记忆单元的哪些信息输出到下一时间步

$$o_t = \sigma\left(w_o \cdot \left[h_{t-1}, x_t\right] + b_o\right) \tag{3}$$

记忆更新：

候选记忆

$$\tilde{c} = \tanh\left(w_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{4}$$

最终记忆

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c_t} \tag{5}$$

当前输出

$$h_t = o_t \cdot \tanh\left(c_t\right) \tag{6}$$

其中，W 矩阵表示各种门控和记忆单元 cell 的参数矩阵，x 代表输入，h 代表隐藏变量，主要用来存储和更新历史信息数据，σ 代表 sigmoid 激活函数，tanh 代表 tanh 激活函数。LSTM 结构图如图 1 所示：

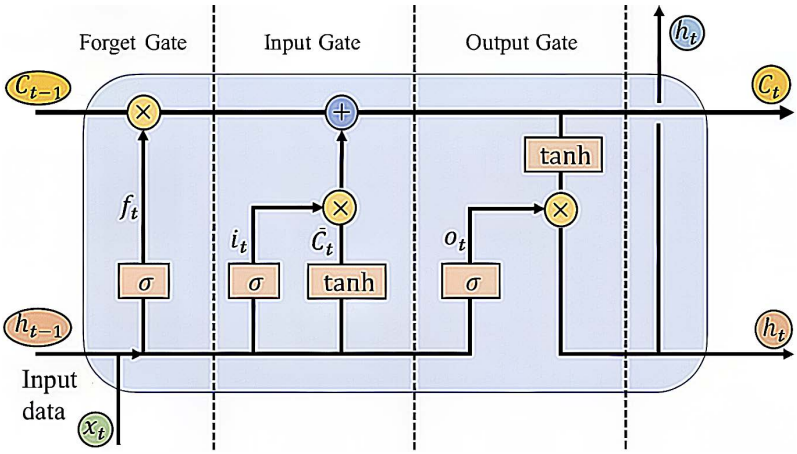

Figure 1. Schematic diagram of LSTM structure
图 1. LSTM 结构原理图

2.4. Transformer 模型

Transformer 模型它是一种基于自注意力机制的架构，拥有并行计算能力，最初设计用于解决机器翻译问题，后来被广泛应用于各种序列建模问题，显示出其广泛的适用性。不同于 LSTM，Transformer 不含有循环连接，而是基于注意力机制构建，这使得模型能够在处理序列数据时摆脱传统的逐步处理限制，实现数据处理的高效并行。Transformer 模型主要由编码器和解码器组成。每个部分都是由多个编码层或者解码层组成。Transformer 模型的主要结构图及其功能图 2 所示：

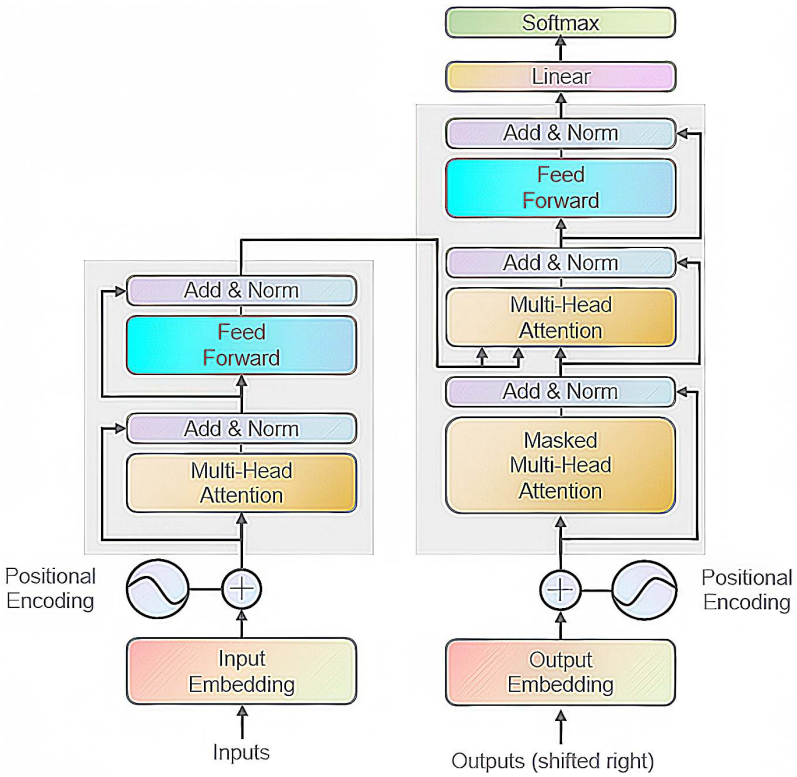

Figure 2. Architecture diagram of the Transformer model
图 2. Transformer 模型架构图

(1) 嵌入层和位置编码层

在训练过程中，原始的时间序列数据无法直接被 Transformer 模型处理，而是需要先经过词嵌入层进行转换，以适应模型所需的特定维度格式。通常 Transformer 模型与采用时间序列逐步传递信息的 RNN 结构不同，它是一次性获取整个数据集，不能保证输入的数据按顺序传递。为了弥补这一点，模型引入了位置编码层，其目的是为序列中的每个数据添加位置信息，确保能够识别和维持这些数据在序列中的相对位置，避免输入的时间序列数据变得无序。常见的位置编码方法是引入正弦函数和余弦函数进行处理，具体的计算公式如下所示：

$$PE\left(pos,2i\right)=\sin\left(\frac{pos}{10000^{2i/d}}\right) \tag{7}$$

$$PE\left(pos,2i+1\right)=\cos\left(\frac{pos}{10000^{2i/d}}\right) \tag{8}$$

其中，pos 表示序列中词语的位置，i 表示位置编码向量的维度索引，d 表示词嵌入向量的维度。

(2) 编码器

编码器模块由多个结构相同但参数不同的编码器层堆叠而成，用于捕捉序列的长期依赖关系。每个编码器包含一个多头注意力机制和全连接前馈神经网络，在这两层之后分别连接一个残差连接块和层归一化块。每个编码器的输出结果都作为下一层编码器的输入，层层传递直至最后一层编码器，最后一层输出的编码信息直接输入解码器中。

通常，自注意力机制与其他深度学习模型相结合使用，以实现更高的预测精度，其计算公式如下所示：

$$Q=XW^{Q},K=XW^{K},V=XW^{V} \tag{9}$$

$$\text{Attention}\left(Q,K,V\right)=\text{Softmax}\left(\frac{QK^{\text{T}}}{\sqrt{d_k}}\right) \tag{10}$$

其中，W^Q，W^K，W^V是可训练的权重矩阵，Q、K、V 分别表示查询矩阵、键矩阵和值矩阵，点积得到的结果除以 $\sqrt{d_k}$ 是为了防止数值过大引起的梯度消失问题。

多头注意力机制由多个并行的自注意力模块组成。每个模块使用独立的权重矩阵，将输入序列分别映射为查询、键和值矩阵，然后计算自注意力。这些模块的输出拼接后通过线性变换合并，保持输入输出维度一致。这种并行结构让模型能同时关注不同位置的特征，更全面地捕捉序列依赖关系，显著提升了特征提取能力。

编码器中引入残差连接块是为了增强模型的稳定性，避免出现梯度消失和梯度爆炸问题。归一化处理的目的是使数据分布更加均匀，从而加速模型的收敛并提升稳定性。

前馈神经网络包含两个全连接层，对输入序列进行两次线性变换处理和一次ReLU非线性变换处理，计算公式如下所示：

$$\text{FFN}\left(x\right)=\max\left(0,xw_1+b\right)\cdot w_2+b_2 \tag{11}$$

式中：w_1 和 w_2 是线性变换层的权重矩阵，b_1 和 b_2 是线性变换层的偏置向量。首先通过第一层线性层将输入特征映射到高维空间，经 ReLU 激活函数进行非线性特征选择；随后通过第二层线性层将特征投影回原始维度，确保输入输出维度的一致性。该模块的输出经过残差连接和层归一化处理后，既保留了原始信息流，又实现了特征的有效转换，最终作为下一编码器层的输入。这种设计在保持模型维度稳定的同时，通过非线性变换增强了特征的表达能力。

(3) 解码器

解码器模块的内部结构与编码器模块类似，但有略微不同，其中，掩码多头自注意力层通过下三角矩阵屏蔽未来信息，确保自回归预测时仅依赖历史数据；编码器-解码器注意力层则通过将编码器最终输出作为 K、V，解码器中间结果作为 Q，实现了输入输出序列的跨模态交互。这种架构设计既保持了 Transformer 并行处理的优势，又满足了时序预测对因果关系的严格要求，有效防止了信息泄露问题。

(4) 线性层和激活函数

解码器最后一层的输出首先通过线性变换层，将学习到的高维特征空间映射至目标维度空间；随后经过非线性激活函数处理，将特征表示转化为概率分布输出。这种“线性变换 + 非线性激活”的级联结构，不仅能够有效建模文本数据中的复杂非线性关系，还能更好地捕获深层次的语义特征表示，从而显著提升模型在语义理解任务中的表现力和泛化能力。

2.5. 模型评价指标

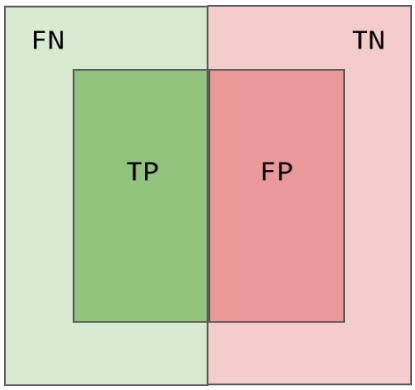

Figure 3. Schematic diagram of basic model evaluation indicators
图 3. 模型评价基本指标示意图

本文分类任务的评价指标主要来自混淆矩阵的相关指标。其中，TP：正类被正确预测为正类的样本数；TN：负类被正确预测为负类的样本数；FP：负类被错误预测为正类的样本数；FN：正类被错误预测为负类的样本数；总样本数为 N = TP + TN + FP + FN。指标直观示意图如图 3 所示。

$$\text{准确率(Accuracy)} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{精确率(Precision)} = \frac{TP}{TP+FP}$$

$$\text{召回率(Recall)} = \frac{TP}{TP+FN}$$

$$\text{F1 分数(F1-Score)} = 2\times\frac{Precision\times Recall}{Precision+Recall}$$

3. 数据预处理及特征选择

3.1. 数据合并

在处理缺失值和异常值之前，我们先进行了数据合并。具体来说就是把水质类别为I类的命名为类别 0，水质类别为II类的命名为类别 1，水质类别为III，IV，V，劣V类的合并为类别 2，一方面是因为I，II类水质数据占了总数据的 90%以上，IV，V，劣V类水质只占了 10%以下，甚至有些观测站不存在这一类

水质数据；另一方面是为了避免出现在不同类别里面的样本量极度不平衡的情况出现，这种情况会对分类效果产生很大的影响。

3.2. 缺失值处理

由于水质数据经常会由于站点维修，设备损毁，人为记录失误等原因而导致出现缺失值，对于这种情况，本文采用的是 k 近邻法进行缺失值的填补，其核心思想是通过寻找与缺失样本相似的非缺失样本，用这些样本的特征值来填补缺失值。详细步骤如下：(1) 对非缺失数据特征进行标准化处理；(2) 计算缺失样本与其他样本的距离，本文选取的是欧氏距离；(3) 选择距离最近的 k 个邻居，本文选取 k = 5；(4) 用这个 k 个邻居的特征值填补缺失值，本文用的是均值。

3.3. 异常值处理

水质数据的异常值和其他数据有一些不同，主要有两种原因，一种是由于自然原因，比如暴雨，气温上升等，或者是人为排污等原因都会导致水质的异常值的产生；还有一种是记录错误，也就是说水质数据的异常值分为真实数据和错误数据，所以我们对于异常值的处理的核心思路是“先识别再针对性处理”，优先保留真实数据、剔除或修正错误数据。思维流程图如图 4 所示：

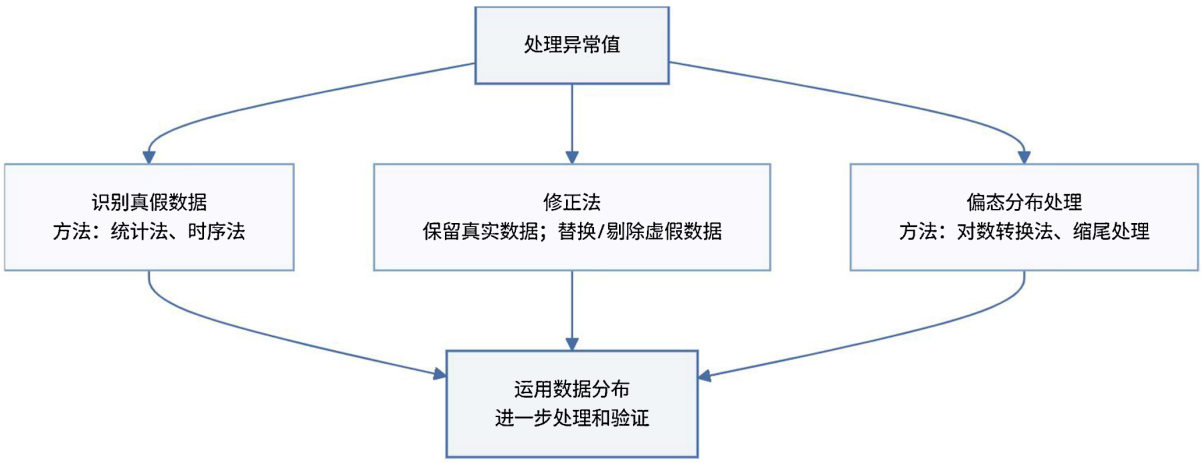

Figure 4. Flowchart of outlier processing ideas
图 4. 异常值处理思路图

统计法为用箱线图的 1.5 倍四分位距、Z-score，|Z| > 3 时识别单指标异常时序法：时序法为用移动平均±2 倍标准差识别突变点。修正法：若异常因传感器故障等，用同时间段邻近监测点数据插值替换；剔除法：仅剔除“孤立异常点”，即周围数据正常，仅单点异常的情况，且剔除比例 ≤5%，避免丢失过多数据；替换法：异常比例较高时，用“中位数替换”和“时序插值”法，来适配时序监测数据。对数转换法：适用于右偏分布的指标，降低极端值影响；winsorization (缩尾处理)：将超出 95%分位数的异常值，替换为分位数数值，比如 95%分位数为 30，将 30 以上值改为 30，这样不会丢失样本。

3.4. 特征选择

为了提高模型的预测精度，以及寻找哪些指标和水质类别的关联程度更高，本文采用随机森林法对特征重要性进行排序。图 5 展示了各指标与水质类别的关联程度。

由图可知，总磷(0.511)是对模型影响最大的核心特征；氨氮(0.159)、浊度(0.109)、高锰酸盐指数(0.108)

的重要性次之；溶解氧、pH 等指标的重要性极低(≤0.044)，对模型贡献较小。而总磷 + 氨氮 + 浊度 + 高锰酸盐指数这四个指标的累计重要性达到了 88.7%，已覆盖核心信息，所以我们优先选择这 4 个特征。

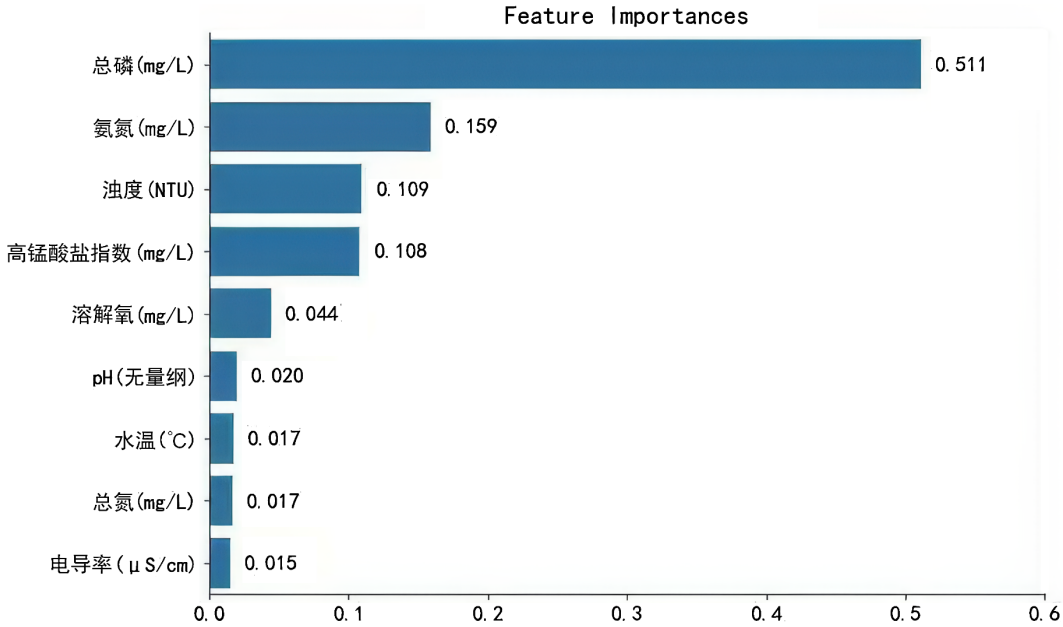

Figure 5. Correlation degree diagram between each indicator and water quality grade
图 5. 各指标与水质类别的关联程度图

4. 融合模型的机制和实证实验

4.1. 融合模型理论

CNN-LSTM-Transformer 融合模型的理论基础建立在特征空间互补性原理之上。这一原理表明，不同深度学习架构能够从相同输入数据中提取出互补的特征表示，通过融合这些特征可以构建更加完整和准确的特征空间。

CNN 的局部特征提取机制基于卷积运算的局部连接特性，通过卷积核在时间序列上的滑动，能够有效捕捉水质指标在局部时间窗口内的关联模式。在水质监测场景中，CNN 主要负责提取同一时间步内不同水质指标。例如，在某一时刻，pH 值的变化往往与溶解氧浓度存在特定的耦合关系，CNN 能够通过卷积操作自动学习这种局部关联模式。

LSTM 的时序依赖建模能力源于其独特的门控机制，包括遗忘门、输入门和输出门。这些门控结构使得 LSTM 能够选择性地记忆和遗忘信息，从而有效处理水质数据中的长期依赖关系。在水质监测中，许多水质指标具有明显的时序特征，如温度的日变化周期、污染物浓度的季节性波动等，LSTM 通过维护细胞状态和隐藏状态，可以捕获这些复杂的时序模式。

Transformer 的全局关联建模优势体现在其自注意力机制上。与传统的序列模型不同，Transformer 能够直接计算序列中任意两个位置之间的关联权重，无需依赖递归或卷积操作。这使得 Transformer 在处理水质数据中的长距离依赖关系时具有显著优势，例如上游污染源对下游水质的影响、季节性气候变化对水质的长期作用等。

基于特征空间互补性原理，本研究提出了两种核心融合策略：串联融合和注意力融合。先对这三个模型提取出来的特征进行串联融合，然后利用注意力机制将不同模型的特征加权整合。

4.2. 串联融合机制

串联融合策略采用递进式特征传递机制，其理论基础是特征的层次化表示。在这种架构中，CNN 首先提取水质数据的局部特征，然后将这些特征传递给 LSTM 进行时序建模，最后由 Transformer 进行全局关联分析。串联融合架构的核心在于如何实现不同模型间的特征传递和维度对齐。在水质时序分类场景中，输入数据通常表示为三维张量 $X \in R^{B\times T\times F}$，其中 B 表示批次大小，T 表示时间步数，F 表示特征维度。

CNN 层的实现采用一维卷积操作，通过卷积核在时间序列上的滑动提取局部特征。卷积核的大小设置为 3，这样既能捕捉相邻时间步的关联，又不会丢失过多的时序信息。卷积层的数学表达为：

$$H_{\text{CNN}} = \text{ReLU}\left(\text{Conv1D}\left(X, K, S\right) + b\right) \tag{12}$$

其中，K 表示卷积核大小，S 表示步长，b 表示偏置向量。

LSTM 层的实现需要考虑输入维度的匹配问题。由于 CNN 输出的特征维度与 LSTM 的输入要求不一致，需要通过线性变换进行维度转换。LSTM 层的实现包括以下步骤：首先，将 CNN 输出的特征张量进行维度调整，从 $[B,T,C]$ 转换为适合 LSTM 输入的格式 $[B,T,H_{\text{LSTM}}]$，这一转换通过全连接层实现：

$$H_{\text{proj}} = \text{Linear}\left(H_{\text{CNN}}, H_{\text{LSTM}}\right) \tag{13}$$

其中 H_{LSTM} 表示 LSTM 的隐藏单元数。然后，将投影后的特征序列输入 LSTM 网络：

$$H_{\text{LSTM}}\left(C_n, h_n\right) = \text{LSTM}\left(H_{\text{proj}}, \left(C_0, h_0\right)\right) \tag{14}$$

其中 C_0, h_0 分别表示细胞的初始状态和隐藏状态。

Transformer 层的实现采用多头自注意力机制，能够直接建模序列中任意两个位置之间的依赖关系。Transformer 编码器的实现包括多个堆叠的编码器层，每个编码器层包含自注意力子层和前馈神经网络子层。在本次水质分类任务中，Transformer 层的输入是 LSTM 输出的隐藏状态序列。通过自注意力机制，模型能够学习到不同时间步和不同特征之间的复杂关联模式。

4.3. 注意力融合机制

注意力融合架构通过引入注意力机制，实现对不同模型特征的动态加权整合。这种融合方式能够根据具体的水质数据特征，自适应地调整各模型输出的重要性权重。注意力机制的设计采用多层感知机(MLP)结构，将 CNN、LSTM、Transformer 的输出特征进行拼接，然后通过非线性变换生成注意力权重：

$$H_{\text{all}} = \text{Concat}\left(H_{\text{CNN}}, H_{\text{LSTM}}, H_{\text{Transformer}}\right) \tag{15}$$

$$a = \text{softmax}\left(\text{MLP}\left(H_{\text{all}}\right)\right) \tag{16}$$

其中，a 是一个三维张量，维度为 $[B,T,3]$，分别对应 CNN、LSTM、Transformer 三种特征的权重。动态权重分配机制的实现需要考虑不同场景下特征重要性的变化。例如，在水质平稳时期，LSTM 提取的时序特征更为重要；而在污染事件发生时，CNN 提取的局部异常特征更关键。通过注意力机制，模型能够自动学习这些场景依赖的权重分配模式。进行权重分配：

$$H_{\text{fusion}} = a_{\text{CNN}} \odot H_{\text{CNN}} + a_{\text{LSTM}} \odot H_{\text{LSTM}} + a_{\text{Transformer}} \odot H_{\text{Transformer}} \tag{17}$$

$$Y = f_{\text{classifier}}\left(H_{\text{fusion}}\right) \tag{18}$$

其中，$\odot$ 表示元素级乘法操作。多层级注意力机制的设计进一步提升了融合效果，除了在模型方面引入注意力机制外，还可以在特征维度和时间维度上分别设计注意力机制，形成多层次的注意力网络，这样设计能够更好地捕捉水质数据在不同尺度上的特征模式。下面给出整体的框架图，见图 6。

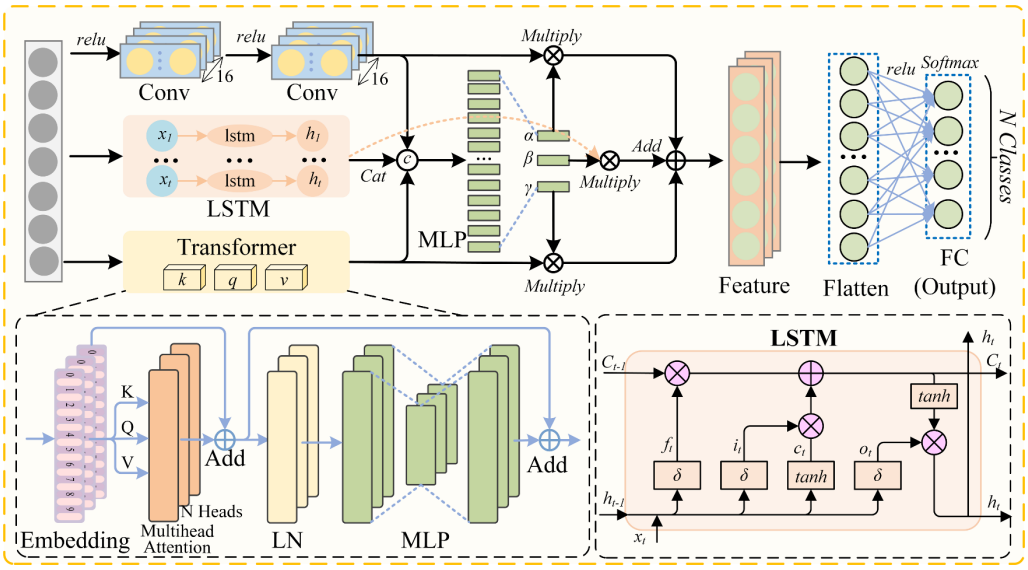

Figure 6. Framework diagram of the fusion model
图 6. 融合模型思路框架图

4.4. 环境配置及参数设置

本文使用 PyTorch 进行编程，用 conda 创建独立环境作为实验环境，优先用 GPU 运行以提高速度。在整个模型训练过程中，使用前 80%的数据作为训练集，后 20%作为测试集，采用了 Adam 优化器，学习率设定为 0.001。模型每个部分的具体参数如表 2 所示：

Table 2. Experimental parameters
表 2. 实验参数表

参数名称	参数含义	参数值
Num-epochs	模型训练的总轮数	400
Batch-size	每次迭代输入模型的样本数	512
Num-classes	水质分类的类别数	3
Window-size	时序滑动窗口大小	4
horizon	预测步长	1
Num-features	每个时间步的水质特征数	4
Input-channels	CNN 的输入通道数	4
Cnn-out-channels	CNN 卷积层的输出通道数	16
Kernel-size	CNN 卷积核大小	3
Hidden-size	LSTM，Transformer 隐层维度	128
Num-layers	LSTM 的层数	2
dropout	随机失活率	0.2
nhead	Transformer 注意力头数	4
Num-transformer-layers	Transformer 编码器层数	2
lr	Adam 优化器的学习率	0.001
Train-size	训练集占总数据集的比例	0.8

4.5. 实验结果

以下两张图反映了 CNN-LSTM-Transformer 融合模型的训练过程，整体表现良好且收敛性稳定。对于准确率曲线，在前 50 个 epoch 内，训练集和验证集的准确率快速提升至 90%以上，最终训练集准确率稳定在 98%左右，测试集准确率稳定在 95%左右，说明模型学习能力强，能快速捕捉水质数据的核心特征。

对于损失曲线，在前 50 个 epoch 内，训练集和验证集的损失快速下降至 0.3 以下，与准确率的快速提升对应。并且没有出现过拟合现象，最后稳定收敛到 0.3 以下。具体结果见图 7，图 8，表 3。

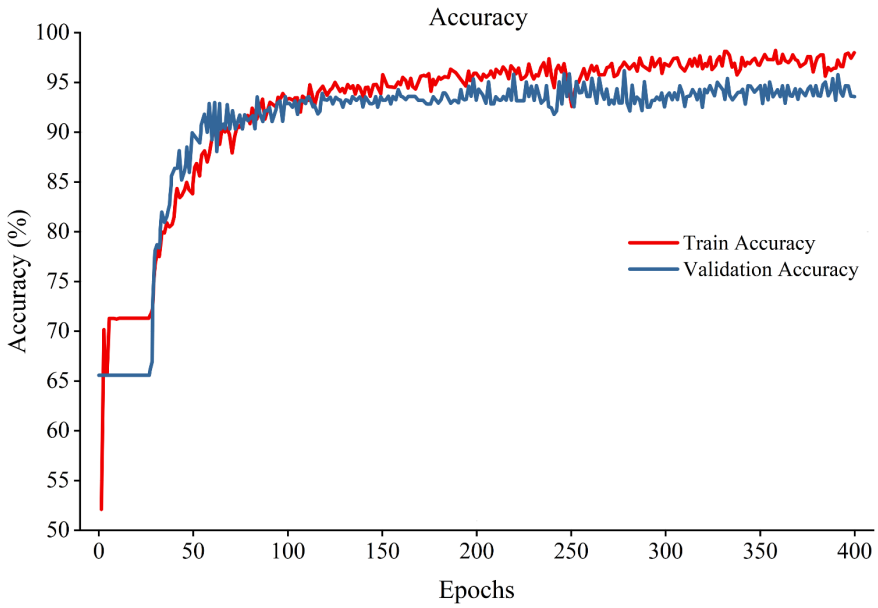

Figure 7. Experimental accuracy results
图 7. 实验准确率结果图

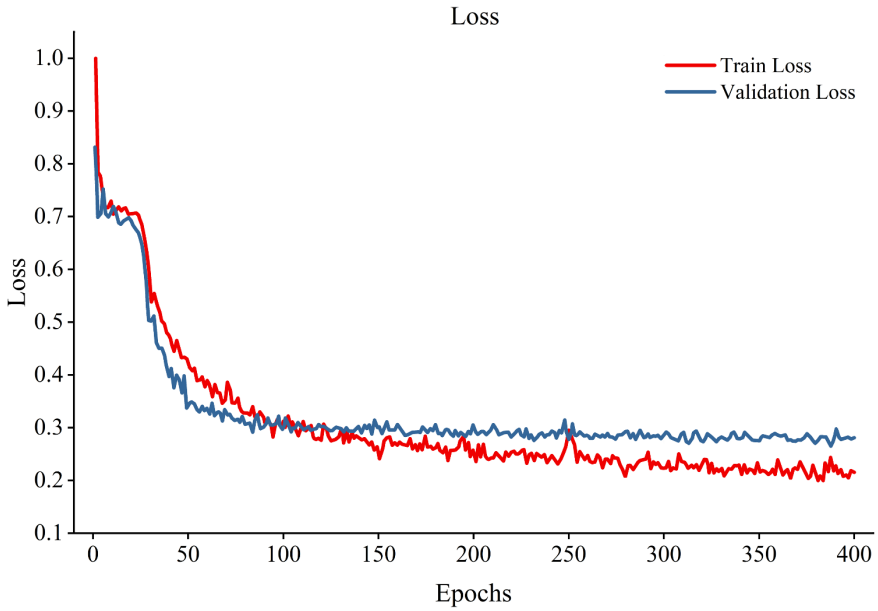

Figure 8. Loss curve of experimental results
图 8. 实验结果损失曲线图

Table 3. Classification results
表 3. 分类结果表

	准确率	损失值
训练集	98.4%	0.19
测试集	96.2%	0.26

5. 消融对比实验和泛化实验

5.1. 消融对比实验

为了验证 CNN-LSTM-Transformer 融合模型的优越性，我们对此模型做了消融对比实验，一方面可以对比单一模型的预测效果，另一方面也可以看到融合模型每个部分的重要性。消融实验的结果图见图 9，具体结果参数见表 4。

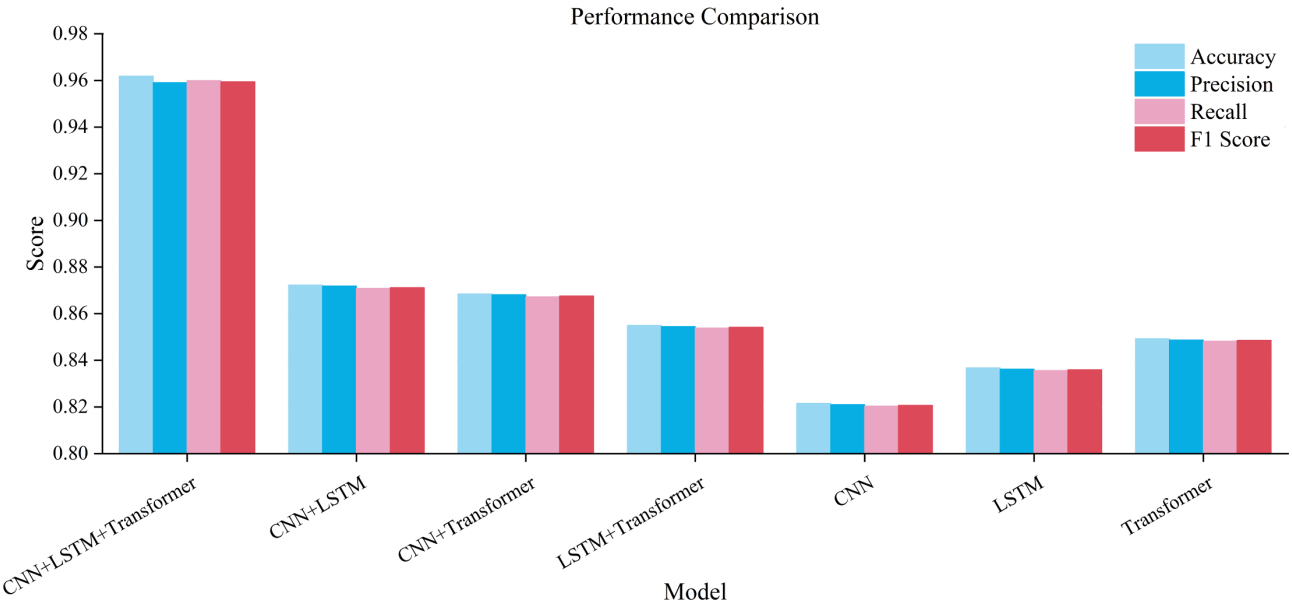

Figure 9. Bar chart of ablation experiment results
图 9. 消融实验结果图

Table 4. Detailed results of ablation experiments
表 4. 消融实验具体结果表

模型名称	Accuracy 准确率	Precision 精确率	Recall 召回率	F1 分数
CNN-LSTM-Transformer	0.962	0.958	0.960	0.959
CNN-LSTM	0.873	0.872	0.870	0.871
CNN-Transformer	0.869	0.868	0.867	0.868
LSTM-Transformer	0.858	0.857	0.856	0.857
CNN	0.823	0.822	0.820	0.822
LSTM	0.838	0.837	0.836	0.837
Transformer	0.829	0.828	0.827	0.828

从结果图 9 中我们可以看到，CNN + LSTM + Transformer 是所有模型中性能最优的，其准确率、精确率、召回率、F1 值均接近 0.96，各项指标在所有模型里处于最高水平。每个模型的准确率、精确率、召回率、F1 值数值都比较接近，说明这些模型在不同评估维度上的表现相对均衡。对于单一模型，CNN 模型的准确率在 82%左右，LSTM 模型准确率在 83%~84%之间，Transformer 模型的准确率在 85%左右，CNN 模型的准确率最低，是因为它只能提取局部特征，很难考虑到全部信息，三个单一模型的预测效果远不如融合模型；而消融后的三个模型，准确率都在 85%~87%之间，虽然预测效果相对单一模型有所提升，但是距离融合模型的效果还有一定差距。

5.2. 泛化实验

为了验证 CNN-LSTM-Transformer 融合模型的泛化能力，避免该模型只对某一数据集的拟合程度高的现象，我们又选取了甘肃省的固水子村和两水桥这两个站点 2023 年的数据，因为前面选取的十个站点都是甘肃省黄河流域的站点，为了验证融合模型的鲁棒性，这次我们选取的数据是甘肃省长江流域的两个站点。实验结果如表 5 所示：

Table 5. Validation results on other datasets
表 5. 其他数据集验证结果表

站点	准确率	损失值
固水子村	0.947	0.21
两水桥	0.935	0.35

从上表可以得知，CNN-LSTM-Transformer 融合模型在其他数据集的分类效果也是很好的，不仅准确率较高，损失值也很低，说明模型的泛化能力强，没有出现过拟合现象。

6. 结论

本文通过构建以注意力机制融合下的 CNN-LSTM-Transformer 模型，对甘肃省选取的 10 个监测站的水质数据进行分类，并且通过消融实验和泛化实验得到以下结论：

(1) 在数据层面：提出差异化异常值处理与随机森林特征筛选方案，更符合水质数据的特点，提升了水质数据质量。选取出了对水质分类影响最大的四个特征，即总磷，氨氮，浊度，高锰酸盐指数这四个特征，减少了工作量。

(2) 在模型层面：构建了注意力机制融合的 CNN-LSTM-Transformer 框架，实现了多维度特征协同提取。

(3) 在实验层面：本文的 CNN-LSTM-Transformer 融合模型相对于单一的机器学习模型有更好的分类效果，其准确率达到了 96.2%，精确率为 95.8%，说明分类效果很好。

基金项目

青海省自然科学基金(No. 2025-ZJ-902T)。

参考文献

[1] 徐祖信. 我国河流单因子水质标识指数评价方法研究[J]. 同济大学学报(自然科学版), 2005(3): 321-325.

[2] 罗芳, 伍国荣, 王冲, 等. 内梅罗污染指数法和单因子评价法在水质评价中的应用[J]. 环境与可持续发展, 2016, 41(5): 87-89.

[3] 邓峰. 大气环境质量综合评价的一种新方法——模糊综合指数法[J]. 干旱环境监测, 1991(2): 118-122+137.

[4] 周林飞, 高云彪, 许士国. 模糊数学在湿地水质评价中应用的研究[J]. 水利水电技术, 2005(1): 35-38.

[5] 侯玉婷, 周忠发, 王历, 等. 基于改进模糊综合评价法的喀斯特山区水质评价研究[J]. 水利水电技术, 2018, 49(7): 129-135.

[6] 谢艾玉, 吕承, 秦梦怡, 等. 变异系数动态赋权-混合隶属度函数耦合的湖泊水质模糊评价[J]. 环境科学学报, 2026, 46(1): 464-474.

[7] 贺北方, 王效宇, 贺晓菊, 等. 基于灰色聚类决策的水质评价方法[J]. 郑州大学学报(工学版), 2002(1): 10-13.

[8] 敖成欢, 钟九生, 赵梦, 等. 基于模糊综合法和灰色关联法的百花湖水质评价[J]. 水土保持通报, 2020, 40(1): 116-122+129.

[9] 葛新越, 王宜怀, 周欣. GA-BP 神经网络在 NB-IoT 水质监测系统中的应用研究[J]. 现代电子技术, 2020, 43(24): 30-33+37.

[10] 方志豪, 李正权, 张铭玮. 基于加权朴素贝叶斯的水质数据分类研究[J]. 物联网学报, 2022, 6(1): 113-122.

[11] 徐玲, 景向楠, 杨英, 等. 基于 SMOTE-GA-CatBoost 算法的全国地表水水质分类评价[J]. 中国环境科学, 2023, 43(7): 3848-3856.

[12] 项新建, 颜超龙, 费正顺, 等. 一种基于多分类器和证据理论融合的水质分类方法[J]. 人民黄河, 2024, 46(1): 109-113.

[13] 刘凯, 刘锋, 彭英湘, 等. 基于机器学习模型的流域水质评价与预测[J]. 环境污染与防治, 2025, 47(7): 1-10.