doi：10.13413/j.cnki.jdxblxb.2023276

基于改进图注意力网络的油井产量预测模型

张 强，彭 骨，薛陈斌
（东北石油大学 计算机与信息技术学院，黑龙江 大庆 163318）

摘要：针对图注意力网络处理噪声和时序数据较弱，并且在堆叠多层后出现梯度爆炸、过平滑等问题，提出一种改进图注意力网络模型. 首先，使用 Squeeze-and-Excitation 模块对样本输入数据的特征信息进行不同程度关注，增强模型处理噪声的能力；其次，使用多头注意力机制，将序列数据中每个序列相对其他序列进行加权求和，提取数据的时序性；再次，将图注意力网络提取的节点特征与节点的度中心性拼接，获取节点的局部特征，并用全局平均池化的方式提取节点的全局特征；最后，将两者进行融合得到节点的最终特征表示，增强模型的表征能力. 为验证改进图注意力网络的有效性，将改进图注意力网络模型与 LSTM，GRU 和 GGNN 模型进行对比，实验结果表明，该模型预测效果得到有效提升，具有更高的预测精度.

关键词：图注意力网络；多头注意力；节点度中心性；全局平均池化
中图分类号：TP18 **文献标志码：**A **文章编号：**1671-5489(2024)04-0933-10

Oil Well Production Prediction Model Based on Improved Graph Attention Network

ZHANG Qiang, PENG Gu, XUE Chenbin
(*School of Computer and Information Technology, Northeast Petroleum University, Daqing* 163318, *Heilongjiang Province, China*)

Abstract: Aiming at the problems that graph attention networks were weak in handling noisy and temporal data, as well as gradient explosion and oversmoothing after stacking multiple layers, we proposed an improved graph attention network model. Firstly, we used the Squeeze-and-Excitation module to pay different levels of attention to the feature information of the sample input data to enhance the model's ability to handle noise. Secondly, the temporal sequence of the data was extracted by using the multi-head attention mechanism, which weighted and summed each sequence in the sequence data relative to the other sequences. Thirdly, the node features extracted from the graph attention network were spliced with the degree centrality of the nodes to obtain the local features of the nodes, and the global features of the nodes were extracted by using global average pooling. Finally, the two were fused to obtain the final feature representation of the nodes, which enhanced the representational ability of the model. In order to verify the effectiveness of the improved graph attention network, the improved graph attention network model was compared with LSTM, GRU and

收稿日期：2023-06-29.

第一作者简介：张 强(1982—)，男，汉族，博士，教授，从事智能进化和神经网络的研究，E-mail：dqpi_zq@163.com. **通信作者简介：**彭 骨(1998—)，男，穿青人，硕士研究生，从事人工智能和神经网络的研究，E-mail：pg24095@163.com.

基金项目：国家自然科学基金(批准号：42002138)、黑龙江省自然科学基金(批准号：LH2022F008)、黑龙江省博士后专项基金(批准号：LBH-Q20077)和黑龙江省优秀青年教师基础研究支持计划项目(批准号：YQJH2023073).

GGNN models. The experimental results show that the prediction effect of the model has been effectively improved, with higher prediction accuracy.

Keywords: graph attention network; multi-head attention; node degree centrality; global average pooling

油井产量是油田开发的重要指标[1]，准确预测产油量可优化开采方案、提高油田开发效率并降低开采成本. 近年来，随着人工神经网络理论的不断发展，越来越多的新型神经网络被应用在油田产量预测中[2-4]. 同时，基于深度学习的神经网络模型也被应用于油田产量的预测. 例如：任燕龙等[5]使用长短期记忆网络建立了油田产量的预测模型；潘少伟等[6]在长短期记忆网络的基础上使用注意力机制提取了数据的时序性，增强了模型的可移植性；张蕾等[7]提出了一种基于时域卷积网络的方法，分阶段搭建产量预测模型，并对该模型使用麻雀搜索算法优化超参数，提高了模型的训练速度. 这些模型虽然都取得了较好的效果，但都未考虑油田开采中井位布局的影响，即未考虑油田井网拓扑结构对产量的影响.

在水驱油田开发过程中，由于油田地质结构复杂和地层非均匀的特性，油田的井位布局采用节点型布局，即油井和水井成交叉或分叉的布置方式，这种布置方式是油田能持续增产的关键因素. 因此，为有效挖掘并利用油田中油井和水井的关系和信息，且体现出不同水井对油井的不同影响程度，可以将油水井的动静态数据的特征信息和井网的拓扑结构抽象表现为图数据，并采用图注意力网络(graph attention network，GAT)进行处理. 相比传统的机器学习，图注意力网络能考虑到图数据的拓扑结构，捕捉图数据中节点与边以及节点与节点之间的关系，具有更强的表达和泛化能力. 但图注意力网络存在处理噪声能力弱，不能处理时序性数据以及在堆积多层后会出现梯度爆炸、过平滑等问题，为解决这些问题，本文提出一种改进的图注意力网络搭建产量预测模型. 首先，采用SE(squeeze-and-excitation)模块[8]调整输入数据特征的重要性，避免噪声特征数据产生的影响，改善模型的鲁棒性；其次，采用多头注意力机制，使序列数据中每个序列占有不同的权重，使时间轴上越近的序列所占的权重越大，提高图注意网络处理时序数据的能力；最后，采用多尺度特征提取的方法，分别获取图数据中节点的全局特征和局部特征，并将两者相融合得到节点的特征表示，以提高模型的表征能力.

1 图注意力网络模型

图神经网络[9]可使用图模型处理复杂的系统建模问题，其核心思想是通过邻接矩阵的方式聚合节点和邻居节点的特征，得到节点的最终输出表示[10-11]. 但这种聚合方式存在缺点，不能很好地突出邻居节点对中心节点的影响程度. 图注意力网络[12-14]则改变了节点与邻居节点的聚合方式，不再是简单使用邻接矩阵的方式聚合，而是在迭代过程中用注意力机制方式计算每个邻居节点的注意力权重，即突出了不同邻居节点对节点的影响程度. 节点与邻居节点的权重系数计算公式为

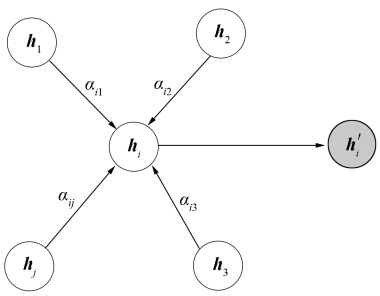


图1 节点聚合

Fig. 1 Node aggregation

$$\alpha_{ij}=\frac{\exp\{\mathrm{LeakyReLU}(\beta^T(\boldsymbol{W}\boldsymbol{h}_i\,\|\,\boldsymbol{W}\boldsymbol{h}_j))\}}{\sum_{k=1}^{N_i}\exp\{\mathrm{LeakyReLU}(\beta^T(\boldsymbol{W}\boldsymbol{h}_i\,\|\,\boldsymbol{W}\boldsymbol{h}_k))\}},\quad(1)$$

其中：$\boldsymbol{h}_i$ 表示中心节点；$\boldsymbol{h}_j$ 表示邻居节点；$\boldsymbol{W}$ 为可训练参数，使用拼接操作将变化后的 $\boldsymbol{h}_i$ 和 $\boldsymbol{h}_j$ 拼接；用 β^T 表示参数点积操作，同时将得到结果进行归一化即可得到权重系数；α_{ij} 为归一化后的结果，表示 $\boldsymbol{h}_j$ 对 $\boldsymbol{h}_i$ 节点的影响程度. 定义 $\boldsymbol{h}_i'$ 为 $\boldsymbol{h}_i$ 节点更新后的输出表示，更新过程如图1所示，计算公式为

$$\boldsymbol{h}_i'=\sigma\Big(\sum_{j=1}^{N_i}\alpha_{ij}\boldsymbol{W}\boldsymbol{h}_j\Big).\quad(2)$$

2　改进图注意力网络预测模型

为提高 GAT 处理噪声数据和时序数据的能力，在网络中引用 SE 模块和多头注意力机制. SE 模块会自适应调整输入数据特征的重要性，避免噪声数据对模型产生影响. 多头注意力机制会计算不同序列之间的权重，然后使序列数据中的每个序列相对其他序列进行加权求和，以此提取输入数据的时序性. 同时为增强 GAT 网络的表征能力，对输入数据的特征采用多尺度融合的方法. 即将经过数据转换模块得到的图数据输入到特征提取模块中，分别提取图数据中节点的局部特征和全局特征；将获取的局部特征和全局特征进行融合，从而得到节点的最终表示. 模型的整体框架如图 2 所示. 模型流程如下：

1）将样本特征张量 $\boldsymbol{X}$ 输入 SE 模块，使 SE 模块学习张量 $\boldsymbol{X}$ 中每个特征的重要性权重，然后将权重与原始张量 $\boldsymbol{X}$ 相乘得到加权后的特征向量输出；

2）将 SE 模块输出的特征向量输入到多头注意力机制中，利用多头注意力机制学习特征向量中每个序列的权重，使每个序列相对其他序列进行加权求和得到新的输出向量；

3）为提高模型的泛化性，将经过多头注意力机制输出后的特征向量输入到数据转换模块，通过数据转换模块将输入数据转换为图结构数据；

4）为增强模型的表征能力，将转换的图数据输入到特征提取模块，分别提取图数据的局部特征和全局特征，再将两者融合，进而得到最终输出张量 $\boldsymbol{X}'$.

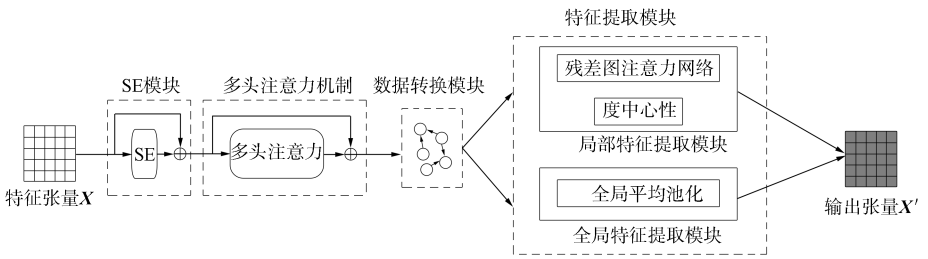


图 2　模型框架

Fig. 2　Model framework

2.1　噪声处理

SE 模块可通过调整每个特征的相关性，抑制不重要的特征，加强模型处理噪声的能力. SE 模块的结构如图 3 所示. 对于给定的张量 $\boldsymbol{X}\in\mathbb{R}^{H\times C}$（其中 H 表示样本数量，C 表示特征数），计算公式为

$$z_i=\frac{1}{H}\sum_{j=1}^{H}x_{i,j}, \tag{3}$$

其中 z_i 表示第 i 个特征多个样本取均值后的结果，图中激励操作是将 z_i 输入两层全连接中，计算出每个特征的注意力权重. 定义 s_i 为张量 $\boldsymbol{X}$ 中第 i 个特征的注意力权重，计算公式为

$$f(z_i)=\mathrm{ReLU}(\boldsymbol{W}_1z_i+b_1), \tag{4}$$

$$s_i=\sigma(\boldsymbol{W}_2f(z_i)+b_2). \tag{5}$$

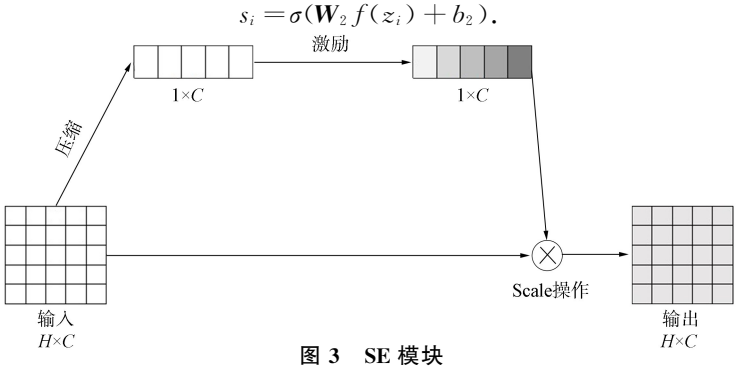


图 3　SE 模块

Fig. 3　SE module

2.2 时序特征提取

注意力机制是一种计算机视觉和自然语言处理等领域常用的技术[15-18]，它可以在给定的输入中自动找出有用的信息，进而提高模型对重要信息的关注度. 其基本思想是为每个输入元素分配一个权重，然后将权重与原始输入相乘得到加权后的输出. 虽然注意力机制在很多任务上性能优异，但注意力机制仍存在计算开销大和不能捕捉长距离依赖关系等缺点. 自注意力机制[19]则克服了这些缺点，是注意力机制的一种特殊形式，它采用并行计算的方式提高模型的计算效率，并且能捕捉输入序列之间的长距离依赖关系. 在自注意力机制中，每个向量都会相对其他向量计算注意力权重，然后进行加权求和获得新的输出向量. 图4为时序数据中单个序列的自注意力机制计算过程.

将图4中序列 x_1 与权重矩阵 $\boldsymbol{W}^Q,\boldsymbol{W}^K,\boldsymbol{W}^V$ 分别相乘即可得到 $\boldsymbol{q}_1,\boldsymbol{k}_1,\boldsymbol{v}_1$，将数据中其他序列与权重矩阵 $\boldsymbol{W}^Q,\boldsymbol{W}^K,\boldsymbol{W}^V$ 分别相乘也会得到对应的 $\boldsymbol{q},\boldsymbol{k},\boldsymbol{v}$. 最终将这些结果相结合即可得到3个矩阵 $\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}$. 利用得到的 $\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}$ 即可计算出经过自注意力机制后的输出表示. 整体计算过程如图5所示，其中 $\sqrt{d_k}$ 是标准化操作.

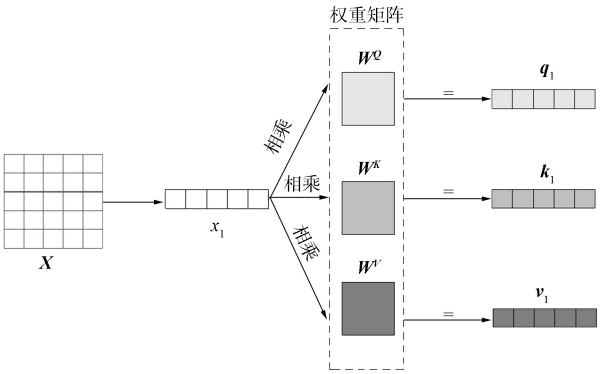


图4 自注意力机制计算过程

Fig. 4 Calculation process of self-attention mechanism

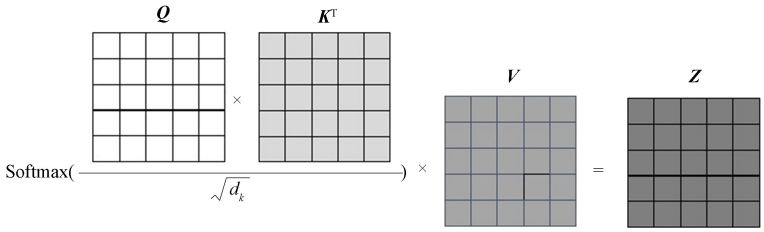


图5 矩阵计算过程

Fig. 5 Matrix calculation process

多头注意力机制则是在自注意力机制基础上，将输入向量映射到不同的子空间，在每个子空间上进行自注意力机制的计算；最后将多个子空间的注意力表示结果合并得到最终的输出表示[20]. 因此，相比自注意力机制，多头注意力机制能捕捉输入的不同方面特征，具有更好的表征能力. 给定一个输入张量 $\boldsymbol{X}\in\mathbb{R}^{N\times d}$，多头注意力机制包含 h 个头，对于自注意力头 i，Head_i 执行如下运算：

$$\mathrm{Head}_i=\mathrm{Softmax}\left(\frac{\boldsymbol{Q}_i(\boldsymbol{K}_i)^{\mathrm{T}}}{\sqrt{d_{\mathrm{Head}}}}\right)\boldsymbol{V}_i, \tag{6}$$

$$\boldsymbol{Q}_i=\boldsymbol{X}\boldsymbol{W}_i^Q,\quad \boldsymbol{K}_i=\boldsymbol{X}\boldsymbol{W}_i^K,\quad \boldsymbol{V}_i=\boldsymbol{X}\boldsymbol{W}_i^V, \tag{7}$$

其中：$\boldsymbol{W}_i^Q,\boldsymbol{W}_i^K,\boldsymbol{W}_i^V\in\mathbb{R}^{d\times d_{\mathrm{Head}}}$ 为自注意头 i 的权重矩阵；d_{Head} 为每个头的维度，$d_{\mathrm{Head}}=\dfrac{d}{h}$. 最后沿维度拼接各个注意力头的输出：

$$\mathrm{MultiHead}(x)=\mathrm{Concat}(\mathrm{Head}_i,\cdots,\mathrm{Head}_h). \tag{8}$$

2.3 模型增强

特征提取模块包含两个模块，分别是全局特征提取模块和局部特征提取模块. 其中全局特征模块

中使用全局平均池化[21-22]提取节点的全局特征. 全局平均池化可对样本的每个特征求平均值，然后再连接成一个向量，以此获取样本的全局信息，计算公式为

$$\text{global_feature}=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{X}_i, \tag{9}$$

其中：N 表示节点总数；$\boldsymbol{X}$ 表示节点的特征矩阵，$\boldsymbol{X}\in\mathbb{R}^{N\times D}$，$D$ 表示每个节点特征向量的维度；$\boldsymbol{X}_i$ 表示节点 i 的特征向量；$\text{global_feature}\in\mathbb{R}^{D}$ 表示节点的全局特征，也可视为整个图的全局特征表示.

在局部特征提取模块中使用 GAT 网络和节点度中心性相结合的方式提取节点的局部特征. 同时，为避免在使用 GAT 网络训练时，堆叠多层出现的模型退化、梯度爆炸等问题，在网络中引用了残差连接[23-24]，如图 6 所示. 图 6 中 $h(x)$ 表示由 GAT 组成的残差块，x 表示前一层网络的输出. 整体过程就是将前一层的输出输入到残差块中得到新的输出，最后再将新的输出输入到后续网络中. 残差块可以将模型的冗余层数变为恒等映射，从而避免产生模型退化现象.

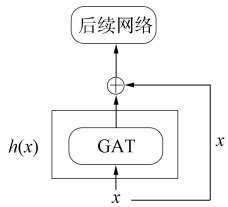


图 6　残差图注意力网络

Fig. 6　Residual graph attention network

为提高模型的表征能力，将经由 GAT 网络后的输出与节点度中心性[25-27]进行拼接，以此表征节点的局部特征. 节点度中心性表示该节点在整个图中的重要程度. 度中心性较高的节点表示它的邻居节点更多，同时也表示该节点的传播能力较强. 如果删除该节点，则会对整个图的连通性造成较大影响. 因此，计算图中节点的度中心性非常必要. 本文主要针对有向图的处理，有向图的节点度中心性需要分别计算入度中心性和出度中心性，然后再将两者相加，计算公式为

$$C_{\text{in}}(i)=\frac{d_{\text{in}}(i)}{n-1}, \tag{10}$$

$$C_{\text{out}}(i)=\frac{d_{\text{out}}(i)}{n-1}, \tag{11}$$

$$C(i)=C_{\text{in}}(i)+C_{\text{out}}(i), \tag{12}$$

其中 $C_{\text{in}}(i)$ 表示节点 i 的入度中心性，$d_{\text{in}}(i)$ 表示节点 i 的入度数，n 表示节点总数. 节点 i 的出度中心性 $C_{\text{out}}(i)$ 与入度中心性计算过程相同. 最后将入度中心性与出度中心性相加即得出节点 i 的度中心性 $C(i)$.

3　算法步骤

下面给出改进图注意力网络模型(SMF-GATs)算法.

算法 1　SMF-GATs 模型.

输入：采油井和注入井每月生产信息 $\boldsymbol{X}=(x_1,x_2,\cdots,x_N)$，采油井和注入井的连通信息 $\boldsymbol{A}=(\alpha_1,\alpha_2,\cdots,\alpha_N)$；

输出：t 时刻采油井的产量 $\boldsymbol{X}_t$；

步骤 1) $\boldsymbol{W}_t\leftarrow\text{fc}(\text{globalAvgPool}(\boldsymbol{X}))$

步骤 2) $\boldsymbol{X}\leftarrow\text{cal}(\boldsymbol{X},\boldsymbol{W}_t)$

步骤 3) for $i\leftarrow 1$ to num do //num 表示 Head 数

步骤 4)　$\boldsymbol{W}_i^{\boldsymbol{Q}},\boldsymbol{W}_i^{\boldsymbol{K}},\boldsymbol{W}_i^{\boldsymbol{V}}\leftarrow\text{linear}(\boldsymbol{X})$

步骤 5)　$\boldsymbol{Q}_i\leftarrow\boldsymbol{X}\boldsymbol{W}_i^{\boldsymbol{Q}}$，$\boldsymbol{K}_i\leftarrow\boldsymbol{X}\boldsymbol{W}_i^{\boldsymbol{K}}$，$\boldsymbol{V}_i\leftarrow\boldsymbol{X}\boldsymbol{W}_i^{\boldsymbol{V}}$

步骤 6)　$\text{Head}_i\leftarrow\text{Softmax}\left(\dfrac{\boldsymbol{Q}_i(\boldsymbol{K}_i)^{\mathrm{T}}}{\sqrt{d_{\text{Head}}}}\right)\boldsymbol{V}_i$

步骤 7) end for

步骤 8) $\boldsymbol{X}_{\text{att}} \leftarrow \text{concat}(\text{Head}_i, \cdots, \text{Head}_{\text{num}})$

步骤 9) for $i \leftarrow 1$ to N do //N 表示月份数

步骤 10) $g_i \leftarrow \text{Graph}(x_i, \alpha_i)$

步骤 11) end for

步骤 12) for $\boldsymbol{G} \leftarrow (g_1, g_2, \cdots, g_N)$, $\boldsymbol{X}_{\text{global}} \leftarrow \text{globalfeature}(\boldsymbol{G})$

步骤 13) $\boldsymbol{X}_{\text{local}} \leftarrow \text{concat}(\text{resGAT}(\boldsymbol{G}), \text{degree}C(\boldsymbol{G}))$

步骤 14) $\boldsymbol{X}_t \leftarrow \text{fc}(\boldsymbol{W}_1 \boldsymbol{X}_{\text{global}} + \boldsymbol{W}_2 \boldsymbol{X}_{\text{local}})$

步骤 15) end for

步骤 16) return $\boldsymbol{X}_t$.

算法 1 的步骤如下：

1) 输入生产信息 $\boldsymbol{X}=(x_1, x_2, \cdots, x_N)$后，算法对 $\boldsymbol{X}$ 中的数据进行降噪处理，即先对 $\boldsymbol{X}$ 进行平均池化压缩得到权重矩阵 $\boldsymbol{W}_t$，再与原始输入信息 $\boldsymbol{X}$ 相乘得到新的输入；

2) 计算降噪输出后 $\boldsymbol{X}$ 的查询($\boldsymbol{Q}$)、键($\boldsymbol{K}$)和值($\boldsymbol{V}$)，并将这三者相乘获得新的输出 $\boldsymbol{X}_{\text{att}}$，以此挖掘出生产信息间的长期依赖关系；

3) 将 $\boldsymbol{X}_{\text{att}}$与采油井和注入井的连通信息 $\boldsymbol{A}=(\alpha_1, \alpha_2, \cdots, \alpha_N)$抽象表现为图数据集 $\boldsymbol{G}$；

4) 在图数据集 $\boldsymbol{G}$ 中分别计算每个节点度中心性和节点聚合周围邻居节点后的新节点，并将两者融合，得到节点的局部特征表示；

5) 采用全局平均池化的方式获取图数据集 $\boldsymbol{G}$ 的全局特征，并将其与前一步的节点局部特征按不同权重融合，得到节点的最终输出表示，再将其输入到全连接中得出采油井在 t 时刻的产油量 $\boldsymbol{X}_t$.

算法 1 是对基础 GAT 模型的改进，涉及降噪、融入注意力机制以及多尺度提取模块等操作，这些操作加强了 GAT 模型处理噪声和时序数据的能力，从而加快了 GAT 模型的训练速度，其时间复杂度低于基础 GAT 模型；但也由于该算法分别计算提取图数据集的局部特征和全局特征，使用了更多的矩阵运算，使得该算法的空间复杂度较基础 GAT 模型高.

4 实验设计与分析

4.1 数据集及归一化处理

本文实验使用的数据集是某油田实际生产过程中产生的油水井月数据，数据集信息列于表 1. 由表 1 可见，油井和水井共有月数据 7 992 条，实验时将数据按月份划分比例. 使用 24 个月的数据作为训练集，3 个月的数据作为测试集，即将所有数据中的 7 104 条数据作为训练集，888 条数据作为测试集进行实验.

表 1 某油田的油水井月数据

Table 1 Monthly data of oil and water wells of certain oilfield

类别	个数	特征数	月份数
油井	170	20	27
水井	126	11	27

实验前需对数据集进行归一化处理. 本文采用最小最大归一化方式，其基本思想就是将原始数据映射到[0,1]内，计算公式为

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{13}$$

其中 x 表示样本中某个特征的值，$x_{\min}$表示该特征在数据集中的最小值，$x_{\max}$表示该特征在数据集中的最大值. 使用最小最大归一化可消除每个特征的量纲影响，有利于模型学习特征之间的隐藏关系.

4.2 数据转换

本文提出的改进图注意力网络预测模型中包含数据转换模块，其主要功能是将模型输入的数据转换为模型内部需要处理的图数据. 对应油田数据集，就是将油水井的月生产数据转换为图数据，方法

如下：将油井和水井视为图数据中的节点，注入井和采油井存在有向边．每个井的特征数据表现为图数据节点的特征．将井距信息作为节点与节点之间边的属性．表 2 列出了某油田月数据中其中 5 个月份的数据．

表 2　某油田油水井图数据

Table 2　Oil and water well graph data of certain oilfield

月份	节点数	特征数	边数
1	296	21	469
2	296	21	469
3	296	21	456
4	296	21	469
5	296	21	465

由表 2 可见，节点之间的边数不是一成不变的，这是因为在油田实际开发过程中可能会存在措施或是调整等生产工况因素，如注水井停用则表现为该注水井与其他采油井并不连通，对应的边数就会减少．

4.3　评价指标

本文实验采用的评价指标有平方绝对误差(MAE)、均方误差(MSE)和 R^2，计算公式如下：

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i|, \tag{14}$$

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2, \tag{15}$$

$$R^2(y,\hat{y})=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}, \tag{16}$$

其中 y_i 为第 i 个样本的真实值，$\hat{y}_i$ 为第 i 个样本的预测值，$\overline{y}$ 为样本真实值的均值，n 为样本的数量．表 3 列出了改进图注意力网络模型在训练集和测试集上的各项指标，MAE 和 MSE 是模型在数据集上的平均误差和均方误差，R^2 体现了模型预测值与真实值的拟合程度．

表 3　改进图注意力网络模型评价指标

Table 3　Evaluation indicators of improved graph attention network model

训练集			测试集集		
MAE	MSE	R^2	MAE	MSE	R^2
0.025 0	0.002 6	0.897	0.019 0	0.001 6	0.878

4.4　实验参数设置

本文实验中，设模型的特征输入维度为 21，多头注意力机制中头数设为 3，局部特征提取模块中 GAT 的层数设为 3．模型整体使用 Adam 优化器作为模型的优化器，加快模型的训练速度．模型的初始学习率为 0.002，训练周期数设为 800，使用均方误差作为模型的损失函数．

4.5　模型分部对比

为验证 SE 模块、多头注意力机制和特征提取模块对图注意力网络模型的提升效果，在相同的训练集和测试集情况下，使用同样的初始学习率、优化器、训练周期等超参数，对上述各部分进行实验对比验证，实验对比结果列于表 4．表 4 中的 S-GATs 表示 SE 模块与 GATs 网络的融合，通过 SE 模块自适应调整输入数据各特征的权重，避免了噪声数据的影响，使基础 GATs 网络拟合效果得到提升；SM-GATs 表示在 S-GATs 网络的基础上融合了多头注意力机制，挖掘出输入数据的时序特征，提升了模型网络的性能；SMF-GATs 表示在 SM-GATs 的基础上加入特征提取模块，提取图数据中的节点局部特征和全局特征，并将两者进行融合，增强了模型的表征能力，加快了模型的训练速度，使模型的预测效果得到提升．

表4 改进图注意力网络不同模块的性能对比

Table 4 Performance comparison of different modules for improved graph attention network

模型	训练集			测试集		
	MAE	MSE	R^2	MAE	MSE	R^2
GATs	0.030 3	0.003 7	0.848	0.026 2	0.002 5	0.814
S-GATs	0.030 1	0.003 5	0.860	0.024 8	0.002 2	0.837
SM-GATs	0.027 4	0.003 0	0.879	0.024 4	0.002 0	0.848
SMF-GATs	0.025 0	0.002 6	0.897	0.019 0	0.001 6	0.878

4.6 多种模型预测性能对比

为进一步验证SMF-GATs模型的有效性，将该模型与长短时记忆网络(long short-term memory，LSTM)[28-29]、门控循环单元(gated recurrent unit，GRU)[30-31]和门控图神经网络(gated graph sequence neural networks，GGNN)[32-33]模型进行对比. 为充分体现各模型的性能，对每个模型的参数进行调优，选择最佳的参数，实验结果列于表5. 由表5可见，SMF-GATs和GGNN模型的R^2值均高于LSTM和GRU模型. 这是因为SMF-GATs和GGNN模型都属于图神经网络，而图神经网络对比其他经典深度学习网络会在训练过程中捕捉数据中节点与节点之间的隐藏关系，从而提高模型的准确性. SMF-GATs模型在训练集和测试集的拟合程度比GGNN模型更高，这是因为SMF-GATs模型引入了SE模块，利用SE模块自适应调整特征的特性，减少了噪声数据对模型的影响，提高了模型的准确性，并且SMF-GATs模型从不同尺度提取数据特征进行融合，增强了模型的自身表征能力.

表5 不同模型预测性能对比

Table 5 Comparison of prediction performance of different models

模型	训练集			测试集		
	MAE	MSE	R^2	MAE	MSE	R^2
LSTM	0.039 8	0.005 8	0.769	0.031 2	0.003 6	0.729
GRU	0.032 4	0.004 4	0.820	0.029 8	0.003 0	0.776
GGNN	0.028 9	0.003 6	0.855	0.024 4	0.002 3	0.824
SMF-GATs	0.025 0	0.002 6	0.897	0.019 0	0.001 6	0.878

综上所述，针对图注意力网络处理噪声和时序数据较弱，并且在堆叠多层后出现梯度爆炸、过平滑等问题，本文提出了一种改进图注意力网络模型. 该模型在数据输入时使用SE模块调整输入数据的特征权重，抑制了数据中不重要的特征，避免噪声数据的影响，改善了模型的鲁棒性. 多头注意力机制的应用则使模型能更好地处理时序数据. 为增强模型的表征能力，在模型中使用残差连接、图注意力网络和节点度中心性相结合方式提取图数据节点中的局部特征，并使用全局平均池化的方式提取节点的全局特征，最后将两者融合得到图数据节点的最终表示，使模型预测准确性得到提升.

参 考 文 献

[1] 计秉玉，孟新. 基于储量结构变化的油田开发指标预测与优化数学模型研究[J]. 运筹与管理，2022，31(9)：135-139.(JI B Y，MENG X. Research on Forecasting and Optimizing Mathematical Model of Oilfield Development Indexes Based on Changes in Reserve Structure[J]. Operations Research and Management Science，2022，31(9)：135-139.)

[2] 叶锋. 多元线性回归在经济技术产量预测中的应用[J]. 中外能源，2015，20(2)：45-48.(YE F. Application of Multiple Linear Regression to Economic and Technical Oil Production Forecast[J]. Sino-Global Energy，2015，20(2)：45-48.)

[3] 钟仪华，林旭旭，刘雨鑫. 基于多元混沌时间序列的油田产量预测模型[J]. 数学的实践与认识，2016，46(6)：99-105.(ZHONG Y H，LIN X X，LIU Y X. The Oilfield Output Prediction Model Based on Multivariate Chaotic Time Series[J]. Mathematics in Practice and Theory，2016，46(6)：99-105.)

[4] 李东昊，王天柱. 基于灰色理论和遗传算法的单井产油量预测方法[J]. 天津科技，2017，44(8)：48-51.(LI D H，WANG T Z. Grey Theory and Genetic Algorithm for Single Well Oil Production Prediction[J]. Tianjin

Science & Technology，2017，44(8)：48-51.）

[5] 任燕龙，谷建伟，崔文富，等. 基于改进果蝇算法和长短期记忆神经网络的油田产量预测模型 [J]. 科学技术与工程，2020，20(18)：7245-7251.（REN Y L，GU J W，CUI W F，et al. Oilfield Production Prediction Model Based on Improved Fruit Fly Algorithm and Long-Short Term Memory Neural Network [J]. Science Technology and Engineering，2020，20(18)：7245-7251.）

[6] 潘少伟，郑泽晨，王吉，等. 基于长短期记忆网络和注意力机制的油井产油量预测 [J]. 科学技术与工程，2021，21(30)：13010-13015.（PAN S W，ZHENG Z C，WANG J，et al. Oil Production Prediction of Oil Wells Based on Long Short-Term Memory Networks and Attention [J]. Science Technology and Engineering，2021，21(30)：13010-13015.）

[7] 张蕾，窦宏恩，王天智，等. 基于集成时域卷积神经网络模型的水驱油田单井产量预测方法 [J]. 石油勘探与开发，2022，49(5)：996-1004.（ZHANG L，DOU H E，WANG T Z，et al. A Production Prediction Method of Single Well in Water Flooding Oilfield Based on Integrated Temporal Convolutional Network Model [J]. Petroleum Exploration and Development，2022，49(5)：996-1004.）

[8] 马帅，刘建伟，左信. 图神经网络综述 [J]. 计算机研究与发展，2022，59(1)：47-80.（MA S，LIU J W，ZUO X. Survey on Graph Neural Network [J]. Journal of Computer Research and Development，2022，59(1)：47-80.）

[9] JIANG B，ZHANG Z Y，LIN D D，et al. Semi-supervised Learning with Graph Learning-Convolutional Networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：2019：11313-11320.

[10] 马超，熊顺，蒋丹妮. 图卷积神经网络在道路网选取中的应用 [J]. 测绘科学，2022，47(12)：200-205.（MA C，XIONG S，JIANG D N. Application of the Graph Convolution Network in the Road Network Auto-selection [J]. Science of Surveying and Mapping，2022，47(12)：200-205.）

[11] VELICKOVIC P，CUCURULL G，CASANOVA A，et al. Graph Attention Networks [EB/OL].（2017-10-30）[2023-03-15]. https://arxiv.org/abs/1710.10903.

[12] 岳子巍，方勇，张磊. 基于图注意力网络的安卓恶意软件检测 [J]. 四川大学学报(自然科学版)，2022，59(5)：88-95.（YUE Z W，FANG Y，ZHANG L. Android Malware Detection Based on Graph Attention Networks [J]. Journal of Sichuan University (Natural Science Edition)，2022，59(5)：88-95.）

[13] 刘渊，赵紫娟，杨凯. 基于节点相似性的图注意力网络表示学习模型 [J]. 计算机应用研究，2023，40(3)：822-827.（LIU Y，ZHAO Z J，YANG K. Graph Attention Network Representation Learning with Node Similarity [J]. Application Research of Computers，2023，40(3)：822-827.）

[14] HU J，SHEN L，SUN G. Squeeze-and-Excitation Networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway，NJ：IEEE，2018：7132-7141.

[15] 牛哲文，余泽远，李波，等. 基于深度门控循环单元神经网络的短期风功率预测模型 [J]. 电力自动化设备，2018，38(5)：36-42.（NIU Z W，YU Z Y，LI B，et al. Short-Term Wind Power Forecasting Model Based on Deep Gated Recurrent Unit Neural Network [J]. Electric Power Automation Equipment，2018，38(5)：36-42.）

[16] CAO P，CHEN Y，LIU K，et al. Adversarial Transfer Learning for Chinese Named Entity Recognition with Self-attention Mechanism [C]//Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing. [S.l.]：ACL，2018：182-192.

[17] 谢琪琦，辛月兰，曾曦. 基于注意力机制的多视图三维重建 [J]. 激光杂志，2023，44(1)：136-142.（XIE Q Q，XIN Y L，ZENG X. Multi-view 3D Reconstruction Based on Attention Mechanism [J]. Laser Journal，2023，44(1)：136-142.）

[18] 付宏建，白宏阳，郭宏伟，等. 融合多注意力机制的光学遥感图像目标检测方法 [J]. 光子学报，2022，51(12)：312-320.（FU H J，BAI H Y，GUO H W，et al. Object Detection Method of Optical Remote Sensing Image with Multi-attention Mechanism [J]. Acta Photonica Sinica，2022，51(12)：312-320.）

[19] VASWANI A，SHAZEER N，PARMAR N，et al. Attention Is All You Need [J]. Advances in Neural Information Processing Systems，2017，30：6000-6010.

[20] 张振坤，张冬梅，李江，等. 基于多头自注意力机制的 LSTM-MH-SA 滑坡位移预测模型研究 [J]. 岩土力学，2022，43(增刊 2)：477-486.（ZHANG Z K，ZHANG D M，LI J，et al. LSTM-MH-SA Landslide Displacement

Prediction Model Based on Multi-head Self-attention Mechanism [J]. Rock and Soil Mechanics, 2022, 43(Suppl 2): 477-486.)

[21] 周飞燕，金林鹏，董军. 卷积神经网络研究综述 [J]. 计算机学报，2017，40(6)：1229-1251.（ZHOU F Y, JIN L P, DONG J. Review of Convolutional Neural Network [J]. Chinese Journal of Computeres, 2017, 40(6): 1229-1251.）

[22] CHEN Z Y, XU Q Q, CONG R M, et al. Global Context-Aware Progressive Aggregation Network for Salient Object Detection [C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020: 10599-10606.

[23] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely Connected Convolutional Networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 4700-4708.

[24] 王进，李颖，蒋晓翠，等. 基于层级残差连接LSTM的命名实体识别 [J]. 江苏大学学报(自然科学版)，2022，43(4)：446-452.（WANG J, LI Y, JIANG X C, et al. Named Entity Recognition of LSTM Based on Hierarchial Residual Connection [J]. Journal of Jiangsu University (Natural Science Edition), 2022, 43(4): 446-452.）

[25] DAS K, SAMANTA S, PAL M. Study on Centrality Measures in Social Networks: A Survey [J]. Social Network Analysis and Mining, 2018, 8: 1-11.

[26] 闫泓任，马国帅，钱宇华. 基于节点度中心性的无监督特征选择 [J]. 数据采集与处理，2019，34(2)：312-321.（YAN H R, MA G S, QIAN Y H. Degree-Centrality Based Feature Selection [J]. Journal of Data Acquisition and Processing, 2019, 34(2): 312-321.）

[27] 黄祖南，郑正喜. 复杂产业网络度中心性研究 [J]. 统计研究，2021，38(5)：147-160.（HUANG Z N, ZHENG Z X. Research on the Degree Centrality of Complex Industry Network [J]. Statistical Research, 2021, 38(5): 147-160.）

[28] SONG X Y, LIU Y T, XUE L, et al. Time-Series Well Performance Prediction Based on Long Short-Term Memory (LSTM) Neural Network Model [J]. Journal of Petroleum Science and Engineering, 2020, 186: 106682-1-106682-11.

[29] 王力，金辉. 基于多头自注意力机制的大气数据系统多故障识别 [J]. 江苏大学学报(自然科学版)，2023，44(4)：444-451.（WANG L, JIN H. Multi-fault Recognition of Air Data System Based on Multi-head Self-attention Mechanism [J]. Journal of Jiangsu University (Natural Science Edition), 2023, 44(4): 444-451.）

[30] 刘栋，杨辉，姬少培，等. 基于多模型加权组合的文本相似度计算模型 [J]. 计算机工程，2023，49(10)：97-104.（LIU D, YANG H, JI S P, et al. Text Similarity Computing Model Based on Weighted Combination of Multiple Models [J]. Computer Engineering, 2023, 49(10): 97-104.）

[31] 胡向东，张婷. 基于时空融合深度学习的工业互联网异常流量检测方法 [J]. 重庆邮电大学学报(自然科学版)，2022，34(6)：1056-1064.（HU X D, ZHANG T. Abnormal Traffic Detection Method for Industrial Internet Based on Deep Learning with Time-Space Fusion [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2022, 34(6): 1056-1064.）

[32] LI Y J, DANIEL T, MARC B, et al. Gated Graph Sequence Neural Networks [EB/OL]. (2015-11-17) [2023-03-01]. https://arxiv.org/abs/1511.05493.

[33] 李盼，解庆，李琳，等. 知识增强的图神经网络序列推荐模型 [J]. 计算机工程，2023，49(2)：70-80.（LI P, XIE Q, LI L, et al. Knowledge-Enhanced Graph Neural Network Model for Sequential Recommendation [J]. Computer Engineering, 2023, 49(2): 70-80.）

（责任编辑：韩 啸）