

GDTN:一种用于生命体征预测的图神经网络^{*}

孙佳琪 马陈悦 张雁皓 李长帅 孟祥源 单慧琳
(无锡学院电子信息工程学院 无锡 214105)

摘要:针对医疗领域中生命体征预测的准确性低、计算量大、性能不佳等问题,提出了一种用于生命体征预测的可变形注意力机制的图神经网络模型。该网络保留了传感器的所有时间采样值,利用全连接层进行编码;使用可变形注意力机制作为消息传递和更新机制,提高了生命体征预测的速度;在解码器方面,采用多头注意力机制,从多尺度、多维度观察和提取信息;将输入的特征复制多份并设置为单独的图节点,提高了可变形注意力的适应性和模型的特征提取能力;采用残差网络作为解码器,替代全连接层。输出层使用 GeLU 激活函数替代了传统的 ReLU 激活函数,解决了激活函数在负半轴信息缺失的问题,有效地提高了预测的精准度。测试结果表明,模型在 P19、P12 和 PAM 3 类数据集上的性能均优于其他模型,各项指标均高于最佳基线性能 2.325%,能够有效预测人体的生命体征。

关键词:生命体征;图神经网络;时间序列;Deformable Transformer

中图分类号: TP183;TN911 **文献标识码:** A **国家标准学科分类代码:** 520.2060

GDTN: A graph neural network for predicting vital signs

Sun Jiaqi Ma Chenyue Zhang Yanhao Li Changshuai Meng Xiangyuan Shan Huilin
(School of Electronics and Information Engineering, Wuxi University, Wuxi 214105, China)

Abstract: This paper proposes a graph neural network with a deformable attention mechanism to address the problems of low accuracy, large computational complexity, and poor performance in predicting vital signs in medical detection. The paper preserves all time-sampled values from sensors and encodes them using a fully connected layer. The deformable attention mechanism is used as the message passing update mechanism in the graph neural network, which speeds up the prediction process. In the decoder, a multi-head attention mechanism is used for feature extraction to allow the network to observe information from multiple scales and dimensions. The input features are copied multiple times and set as separate graph nodes to enhance the adaptability of the deformable attention and the feature extraction ability of the model. Instead of a fully connected layer, a residual network is used as the decoder. In the output layer, the GeLU activation function is used instead of the traditional ReLU activation function effectively improve the accuracy of the prediction by addressing the problem of information loss on the negative half-axis. Experimental results demonstrate that the proposed model achieved high performance on three types of datasets (P19, P12, and PAM), with all performance indicators surpassing those of other models 2.325% and higher than the best baseline performance. This indicates the effectiveness of the proposed model in predicting vital sign.

Keywords: vital signs; graph neural network; time serise; Deformable Transformer

0 引言

近年来,随着人工智能技术的不断成熟,智能医疗技术的发展备受瞩目。如何通过病人历史及当前的生命体

征数据来预测其未来的生命体征,已成为一个技术难点。

人类生命体征预测是基于复杂的生命体征大数据,属于时间序列,深度卷积神经网络模型在生命体征等场景的数据是非欧几里得(non-Euclidean)空间类型,其中多个数

收稿日期:2024-07-09

^{*} 基金项目:国家自然科学基金(62071240,62106111)项目资助

据之间具有复杂的关系和相互依赖,是天然的图数据,除时空图神经网络以外的传统图神经网络(graph neural networks, GNN)算法对于时间序列的处理存在天然的缺陷,而基于长短期记忆人工神经网络(long short-term memory, LSTM)^[1]与门控循环单元^[2]的循环神经网络类型算法虽然可以有效地体现时间性,但是无法体现体征信息所存在的空间联系。2021年, Zhang 等^[3]提出了 Rain-drop 模型并采用 Transformer^[4]注意力机制替代图卷积的信息交互方式,解决了因采样时间存在间隔而难以同时采样的问题,很好地表达了生命体征中潜在的空间关联性,但是由于时间序列存在截断和缺失,无法连续地进行实时预测,密集的体征数据也容易过拟合,且训练需要大量的时间和资源。2022年, Schafer 等^[5]尝试以插值、核函数等方法解决由于采样时间序列不规则而导致预测准确度较差的问题,但是效果依然不佳。Fritz 等^[6]将卷积神经网络(convolutional neural networks, CNN)与 LSTM 结合用于预测手术过后 30 d 的死亡率,兼顾了序列数据中的时间特征,但没有充分考虑到生命体征数据的空间关联性,并且在训练中计算量巨大。

为了解决医疗检测中生命体征预测精确度不高、计算量大、性能不佳等问题,本文提出一种可变形注意力机制的图神经网络(graph deformable transformer network, GDTN)模型用于生命体征预测。在数据预处理方面,将不同长度的时间采样数据编码成相同的维度,以保证时间信息的完整性。在主干网络与解码器中引入了 Deformable Transformer^[7]注意力机制,以提高信息提取及编码能力。将残差网络^[8](ResNet)作为本网络的解码器,以确保解码信息的准确性,进一步提高生命体征预测的精确度。

1 数据集

生命体征的预测从生命体征的已有数据中挖掘统计特征,基于历史和当前的生命体征数据对未来的生命体征进行预测。实验使用的医疗保健和人类活动的数据集有 P19、P12 和 PAM 等。其中 P19^[9]包括 38 803 名患者的数据,由 34 个传感器监测,存储了每个患者未来是否发生败血症的数据。P12^[10]记录了 11 988 名患者在进入 ICU 前 48 h 内的 36 个传感器的时间测量值,样本根据住院时间进行标记。PAM^[11]包含患者由 17 个传感器测量的 8 种日常生命体征数据(脉搏、呼吸、心率、血压、体温、血氧、瞳孔、角膜反射等),共计 5 333 个片段,数据集按 8:1:1 的大小比例被随机洗牌切分为训练集、验证集、测试集。

由于生命体征数据均为不规则时间序列,为了更好地提取特征,将不同长度的时间采样数据编码成相同维度,保证了时间信息的完整性;将固定时间内采样的病人信息输入到预测模型中进行预处理,将输入数据分为两类向量,一类为静态特征向量(例如身高、体重、年龄、性别等),另一类为由传感器采集的动态特征向量(例如心率、血压、

体温等)。

2 网络模型

2.1 网络总体结构

整体网络模型的结构如图 1 所示,由图神经主干网络(graph transformer network, GTN)、可变形注意力编码器(deformable attention encoder)和残差解码器(ResNet decoder)3 部分组成。

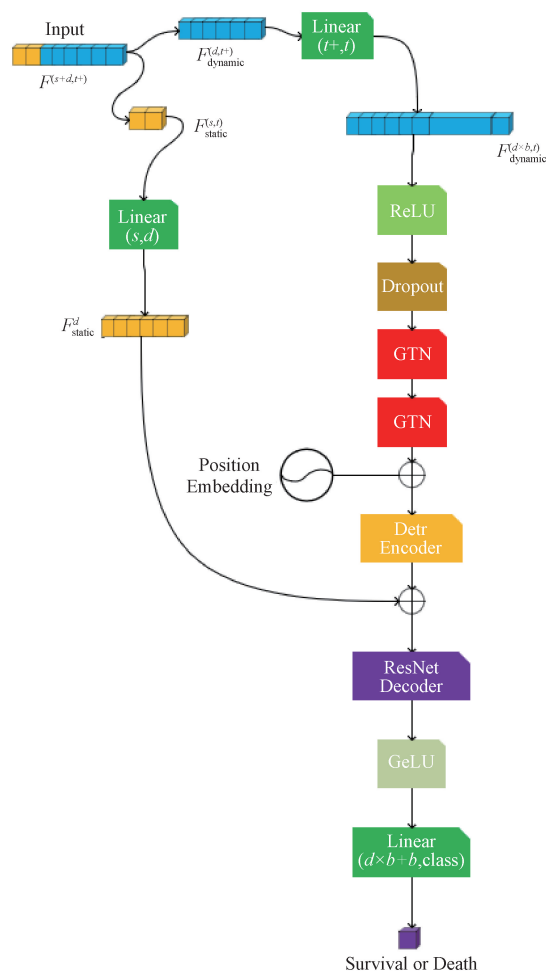


图 1 网络架构

Fig. 1 Network architecture diagram

首先,将固定时间内采样的病人信息输入到预测模型中进行预处理,得到两种类型的向量,即静态特征向量和传感器采集的动态特征向量。将静态特征向量也放入全链接层,同时也将传感器采集的动态特征向量进行复制,采用一层全链接层将所有的时间序列长度由 $t +$ 压缩成 t 。这样有效地保留了时间序列所包含的信息,与传统的插值、核函数和固定截断等方法相比,能避免信息的损失。接着,将特征向量的传感器维度复制并拼接,实现向量信息的复制,添加 ReLU 层和 Dropout 层,能够提高网络模型的正则化效果,提高模型的泛化能力。

其次,将预处理之后的数据输入至 GTN,使用可变形注意力模块结构对相应传感器输入的动态特征向量在每一采样时刻进行图神经展开,建立动态图神经模型以提取图模型的信息。输入和输出均为一个大小不变的二维向量,使用正余弦的位置编码方式,通过此二维向量创建 PE 矩阵,公式如式(1)、(2)所示,标定时间的先后便于进行多维时间序列的编码。

$$PE_{(pos, 2i)} = \sin(pos/1000^{2i/dim}), i \in (0, dim/2) \quad (1)$$

$$PE_{(pos, 2i+1)} = \cos(pos/1000^{(2i+1)/dim}), i \in (0, dim/2) \quad (2)$$

式中: pos 表示位置,从 0 开始编号; dim 表示维度索引。

可变形注意力编码器使用可变形注意力模块融合了来自各传感器的时间序列数据,并对固定时间内的时间信息进行整合。该模块的输入是二维特征向量,输出为三维特征向量。

将上述输出的三维特征向量放入残差解码器进行解码和分析,结合 CNN 处理 GNN 编码信息,从而更有效地利用编码信息。残差解码器的输入为三维特征向量,输出为二维特征向量,将该二维向量与静态特征向量拼接,实现对特征的整合。最后经过 GeLU 激活函数和全连接层,输出人体生命体征的预测信息。

2.2 主干网络

在自然语言处理与计算机视觉等领域经常使用 Attention 注意力机制,如 NAT^[12]、VIT^[13] 和 Swin Transformer^[14]等,此类网络模型能够有效提取图像特征,但存在感受野较小、算力要求高、收敛困难等问题。Zhang 等^[3]将 Transformer 的 Attention 注意力机制与图学习的信息传递更新网络相结合,利用 GNN 自身的 non-Euclidean 动态网状结构,配合 Attention 注意力机制对图数据和图节点进行更新和消息传递。主干网络如图 2 所示,将可变形注意力模块(deformable attention module)结构用于 GTN 的特征提取,对提取的特征进行编码,不仅提高了注意力机制结构对图像特征的提取效果,还减少了模型运算的参量,提高了运算速度。同时,对主干网络输入的特征在第二维度进行复制,增设为单独的图节点,使其符合可变形注意力模块的特征处理方式,提高了输入信息冗余和特征提取的能力。通过多头注意力机制,从多头不同的 GNN 中提取所需特征向量并在第三维度拼接,最终输出维度为 3 的特征向量。

式(3)为输入的二维特征映射即 $\mathbf{x} \in \mathbf{R}^{d \times t}$,将其复制 b 次并拼接成为 $\mathbf{x} \in \mathbf{R}^{(d \times b) \times t}$,创建拥有 $d \times b$ 个节点的动态特征图 $\mathbf{G}, \mathbf{Z}_q$ 作为 query 采样节点的特征向量, $\mathbf{P}_q$ 为所取图的邻接矩阵, $\mathbf{G}$ 为输入的特征图。

$$DeformAttn(\mathbf{Z}_q, \mathbf{P}_q, \mathbf{G}) = \sum_{m=1}^M \mathbf{W}_m \left[\sum_{k=1}^K A_{mqk} \cdot \mathbf{W}'_m \mathbf{x}(\mathbf{P}_q + \Delta \mathbf{P}_{mqk}) \right] \quad (3)$$

式中: m 表示为 Attention 的多头索引; k 为采样图节点邻

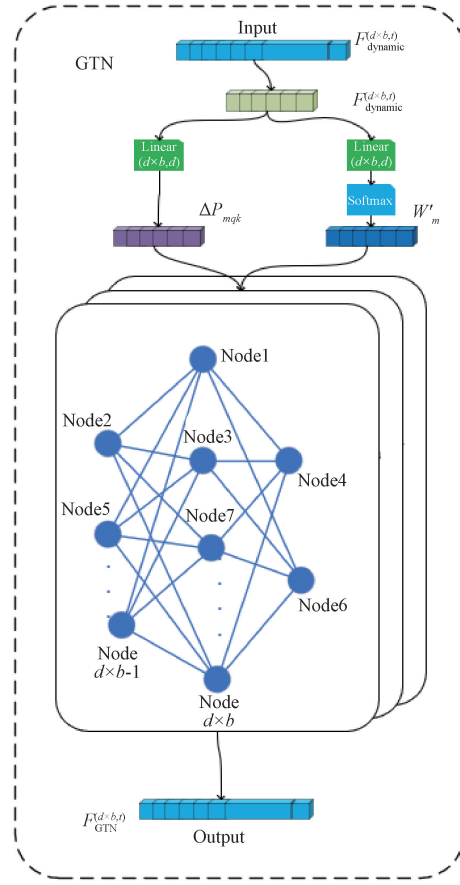


图 2 基于 Deformable Transformer 的图神经更新传播网络

Fig. 2 Graph neural update propagation network based on Deformable Transformer

接矩阵对应邻接图节点的索引; K 为采样点周围特征向量的数值; ΔP_{mqk} 和 A_{mqk} 分别表示第 m 个关注头中第 k 个采样点的采样索引偏移量和注意力权重。每一个关注头的 k 个 A_{mqk} 值的和为 1。而对于 ΔP_{mqk} 存在分数和小数的情况下,对临近的两个索引对应的图节点特征向量的值进行线性插值。

2.3 可变形注意力编码器

基于 Deformable Transformer 的编码器如图 3 所示,采用可变形注意力模块作为编码器,替代了 Raindrop 网络中的 Transformer Encoder。与 Transformer 的注意力模块相比,可变形注意力编码器在显著降低算力需求的同时,通过将二维输入向量复制并扩展为第三维度 c ,采用 Multi-head 多头注意力方式,使得编码器能从多尺度、多维度提取特征信息,编码器的输出在第 3 个特征维度进行扩增,使得编码器输出的向量大小为输入的 N 倍,从而提高了提取信息的能力。

如式(4)所示,对于输入的特征映射 $\mathbf{x} \in \mathbf{R}^{C \times H \times W}$, $\mathbf{Z}_q$ 作为 query 特征向量, $\mathbf{P}_q$ 为所取特征向量在特征图上的二维坐标, $\mathbf{X}$ 为输入的特征图值, m 为 Attention 的多头索

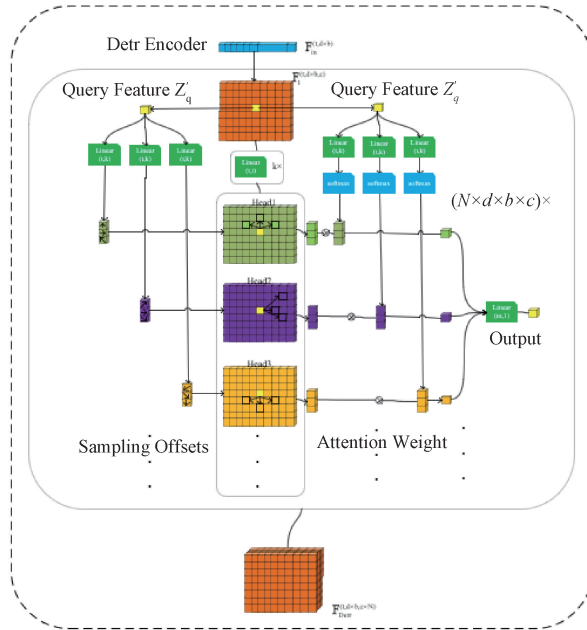


图3 基于 Deformable_Transformer 的编码器网络
Fig. 3 Encoder network diagram based on Deformable_Transformer

引, k 为采样点周围特征向量的索引, K 为采样点周围特征向量的数值, ΔP_{mqk} 和 A_{mqk} 分别表示第 m 个关注头中第 k 个采样点的采样偏移量和注意力权重。每一个关注头的 k 个 A_{mqk} 值的和为 1。

$$\text{DeformAttn}(Z_q, P_q, X) = \sum_{m=1}^M W_m \left[\sum_{k=1}^K A_{mqk} \cdot W'_m x(P_q + \Delta P_{mqk}) \right] \quad (4)$$

2.4 残差解码器

本文使用 CNN 作为解码器,如图 4 所示,通过结构重组和减少权值的方式将特征抽取功能融合进多层感知器,利用参数共享有效地减少了参数的需求量。编码器的输入为三维矩阵特征向量,符合计算机视觉领域的图像特征图。相比之下,传统的 Raindrop 模型直接舍弃了部分三维特征向量,将其转化为二维特征向量,导致其对编码信息的提取不完整,降低了预测结果的准确性。本文使用 CNN 对输入的三维特征向量进行解码提取,并利用残差解码器对 GNN 产生的编码信息进行分析和解读。鉴于数据的特殊性,本文将 (3,3) 大小的卷积核变为 (1,1) 以替代全连接层,对相应特征进行解析。编码器在激活函数选择上使用了 GeLU,能够保留负网络信息且在原点处光滑可导。提高了特征提取地准确性。

2.5 损失函数及优化器

在固定时间内预测病人的体征信息是一个多分类任务,因此本文模型使用交叉熵损失函数 (cross entropy loss) 作为损失函数来实现快速收敛,其表达式如下:

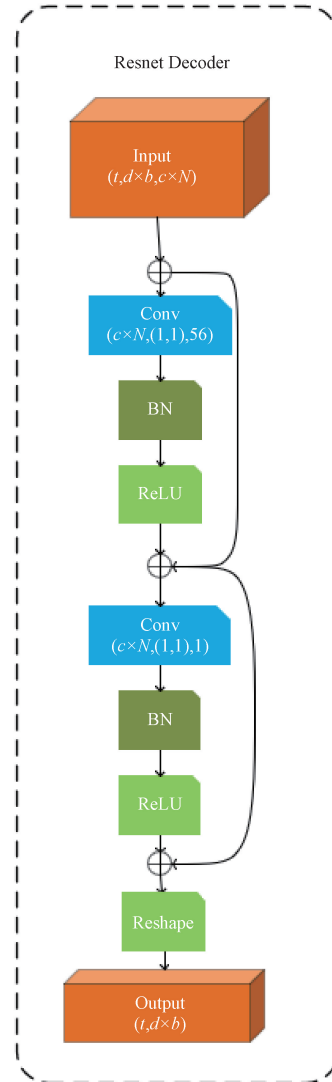


图4 基于 Resnet 的解码器网络
Fig. 4 Resnet-based decoder network diagram

$$\text{Loss} = -\frac{1}{m} \sum_{i=1}^m \sum_{j=1}^N P(x_{ij}) \log(Q(x_{ij})) \quad (5)$$

式中: m 为一个 Batchsize 中的样本个数; N 为分类模型的分类个数; X_{ij} 为分类的物体; $P(x_{ij})$ 为对应分类物体的标签, $Q(x_{ij})$ 为对应分类物体的概率。

使用 AdamW 优化器^[13],它结合了 Adam 与 L2 正则化的优化算法来设计优化器模型。弥补了 Adam 在更新规则中无法收敛到最佳位置的缺陷,加速了损失函数的收敛。AdamW 优化器更新公式如下:

$$\nabla f_t(q_{t-1}) = \text{SelectBatch}(q_{t-1}) \quad (6)$$

$$g_t = \nabla f_t(q_{t-1}) \quad (7)$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (8)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (9)$$

$$\hat{m}_t = m_t / (1 - \beta_1^t) \quad (10)$$

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \quad (11)$$

$$\eta_t = \text{SetScheduleMultiplier}(t) \quad (12)$$

$$\theta_t = \theta_{t-1} - \eta_t (a \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon) + \lambda \theta_{t-1}) \quad (13)$$

式中: t 为更新的步数(steps); a 为学习率; 默认为 0.001, 用于控制步幅(stepsize); θ 为要求解(更新)的参数; $\text{SelectBatch}(q)$ 为选择相应的 Batchsize 获取梯度; η_t 为学习率; $\text{SetScheduleMultiplier}(t)$ 为根据 t 来调整学习率 η_t ; $f(\theta)$ 为目标函数即损失函数; g_t 为目标函数 $f(\theta)$ 对 θ 的求导所得的梯度; β_1 为一阶矩衰减系数, 默认为 0.9; β_2 为二阶矩衰减系数, 默认为 0.99; m_t 为梯度 g_t 的一阶矩, 即 g_t 的期望; v_t 为梯度 g_t 的二阶矩, 即 g_t^2 的期望; $\hat{m}_t$ 为 m_t 的偏置矫正, 考虑到 m_t 在初始情况下为 0 偏置; $\hat{v}_t$ 为 v_t 的偏置矫正, 考虑到 v_t 在初始情况下为 0 偏置; λ 为权重衰减; ϵ 为默认常数, 设置为 10^{-8} , 防止 $\sqrt{\hat{v}_t}$ 为 0。

3 评价指标

由于数据集存在不平衡问题, 本文使用的 6 个图像分割评价指标分别为接受者工作特征曲线下面积(area under a receiver operating characteristic curve, AUROC)、精度-召回率曲线下面积(area under precision-recall curve, AUPRC)、准确率(accuracy)、精确率(precision)、召回率(recall)、F1 分数(F1 Score)。分别定义为:

$$AUROC = \int_0^1 TPR(FPR) d(FPR) \quad (14)$$

$$AUPRC = \int_0^1 Precision(TPR) d(TPR) \quad (15)$$

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (16)$$

$$Precision = TP / (TP + FP) \quad (17)$$

$$Recall = TPR = TP / (TP + FN) \quad (18)$$

$$F1 = 2(Precision \cdot TPR) / (Precision + TPR) \quad (19)$$

$$FPR = FP / (FP + TN) \quad (20)$$

式中: TPR(true positive rate)和 FPR(false positive rate)是在二分类问题中衡量模型性能的重要指标。TPR 也称为召回率, 表示模型正确预测为正类的样本在实际正类样

本中的比例, FP(false positive)是假正类的数量, TN(true negative)是真负类的数量, TP(true positive)表示模型正确预测为正类的样本数量, FN(false negative)表示模型错误预测为负类的样本数量。TPR 和 FPR 通常在 ROC 曲线中使用。ROC 曲线以 FPR 为横轴, TPR 为纵轴, 展示了在不同阈值下模型的性能表现。TPR 和 FPR 的取值范围都在 0~1 之间, 理想情况下, TPR 接近 1 且 FPR 接近 0, 即模型具有较高的召回率和较低的误报率。

4 实验结果与分析

实验仿真平台为 Pycharm, 深度学习框架使用 torch1.11.0+cu113, Python 版本为 3.8.16, 实验所使用的计算机配置为 Intel® Core™i7-13 700 KCPU@5.4 GHz, 12 GB 内存, Nvidia GeForce GTX3060 GPU。

使用交叉熵损失函数作为模型训练的 Loss 函数, 以 AdamW 优化器对模型进行优化, 对训练集、验证集、测试集计算均值和标准差进行归一化, 使得数据集具有相同尺度, 每一个模型的 epoch 均设置为 50 进行迭代; 学习率(learning_rate)预设为 0.001; 优化策略中设置 Loss 停止上升时减小学习率; 学习率因子为 0.1, 即每次学习率降低为原本的 0.1; 忍耐次数(patience), 当网络性能不提升次数为 1 时, 使学习率下降; cooldown 即减小学习率之后恢复正常的 epoch 数为 1。

4.1 典型分类实验

GDTN 在 3 类数据集下的性能评价指标对比如表 1 所示。在 3 个不同数据集上的实验及 6 种基准模型的比较显示, GDTN 在 P19 和 P12 数据集中表现出色, 均优于最强的 Baseline, 在 AUROC 中均高出 1.7%, 与 AUPRC 的数据指标也均高出 0.7%。在多分类的数据集 PAM 中, GDTN 也同样均胜过最强的 Baseline, Accuracy 高出 1%, Precision 提升了 1%, Recall 提高了 0.9%, F1 score 也增加了 0.3%。尽管提升幅度不大, 但考虑到生命体征预测的特殊性, 这仍然代表着相当大的差距。

表 1 GDTN 与 9 个模型在 3 类数据集下的性能评价指标对比

Table 1 Comparison of performance evaluation indexes between GDTN and 9 models under three types of data sets (%)

模型	P19		P12		PAM			
	AUROC	AUPRC	AUROC	AUPRC	Accuracy	Precision	Recall	F1 score
Transformer	83.2±1.3	47.6±3.8	65.1±5.6	95.7±1.6	83.5±1.5	84.8±1.5	86.0±1.2	85.0±1.3
Trans-mean	84.1±1.7	47.4±1.4	66.8±4.2	95.9±1.1	83.7±2.3	84.9±2.6	86.4±2.1	85.1±2.4
GRU-D ^[16]	83.9±1.7	46.9±2.1	67.2±3.6	95.9±2.1	83.3±1.6	84.6±1.2	85.2±1.6	84.8±1.2
SeFT ^[17]	78.7±2.4	31.1±2.8	66.8±0.8	96.2±0.2	67.1±2.2	70.0±2.4	68.2±1.5	68.5±1.8
mTAND ^[18]	80.4±1.3	32.4±1.8	65.3±1.7	96.5±1.2	74.6±4.3	74.3±4.0	79.5±2.8	76.8±3.4
IP-Net	84.6±1.3	38.1±3.7	72.5±2.4	96.7±0.3	74.3±3.8	75.6±2.1	77.9±2.2	77.9±2.2
DGM2-O	86.7±3.4	44.7±11.7	71.2±2.5	96.9±0.4	82.4±2.3	85.2±1.2	83.9±2.3	84.3±1.8
MTGNN	81.9±6.2	39.9±8.9	67.5±3.1	96.4±0.7	83.4±1.9	85.2±1.7	86.1±1.9	85.9±2.4
Raindrop	87.0±2.3	51.8±5.5	72.1±1.3	97.0±0.4	88.5±1.5	89.9±1.5	89.9±0.6	89.8±1.0
GDTN	88.2±1.8	53.7±3.6	74.3±2.1	96.5±1.4	90.4±1.4	90.9±1.8	91.8±0.6	91.2±1.2

4.2 编码器消融对比实验

为了体现出不同编码器对网络性能产生的影响,采用对比实验来比较以下3种情况,未使用编码器、使用传统Transformer Attention 编码器网络以及使用 Deformable Attention 编码器。实验选用 P19 数据集,消融实验指标如表2所示,未添加数据噪声,设置 epoch 为 40, Batchsize 为 128。图5所示为训练集与验证集的 Loss 值,可见,采用 Deformable Transformer 与 Transformer 编码器后, Loss 性能得到轻微提升,相较于未使用编码器的模型,损失值降低了 9.7%。图6所示为训练集与验证集的 AUROC 值,结果表明,Deformable Transformer 的性能超越了未使用编码器和 Transformer 编码器,其 AUROC 值分别提高了 3.2% 以及 0.8%。图7所示为训练集与验证集 AUPRC 值,可见,使用 Deformable Transformer 后,其性能显著优于未使用编码器和使用 Transformer 编码器,其 AUPRC 值分别提高了 14.0% 以及 3.1%。

表 2 P19 数据集编码器消融实验数据指标
Table 2 P19 dataset encoder ablation experimental data indicators

模型	Loss	AUROC/%	AUPRC/%
无编码器	0.611	83.6±2.2	44.4±4.9
Transformer	0.559 8	85.6±2.5	49.1±3.8
Deformable Transformer	0.551 7	86.3±1.8	50.6±3.4

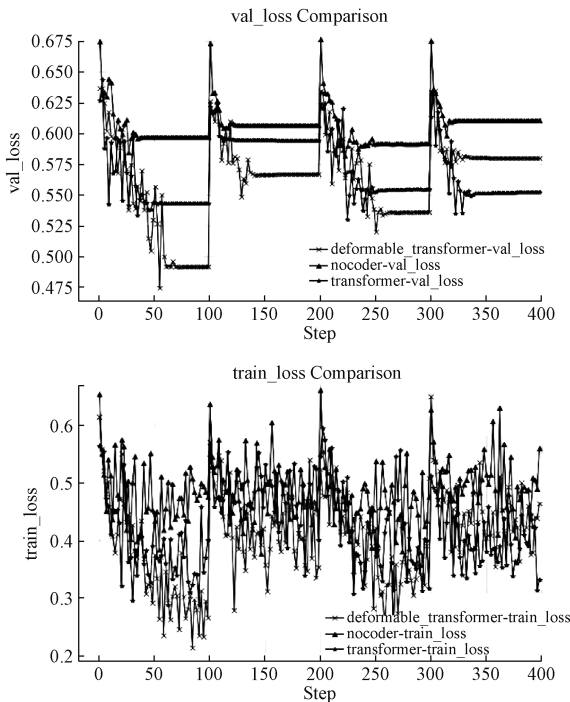


图 5 不同编码器训练集与验证集的 Loss 值
Fig. 5 Loss values of different encoder training sets and verification sets

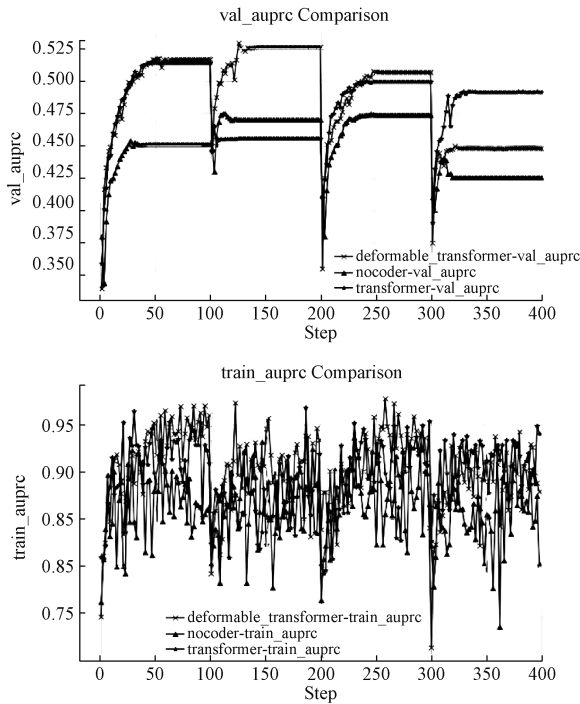


图 6 不同编码器训练集与验证集的 AUROC 值
Fig. 6 AUROC values of different encoder training sets and validation sets

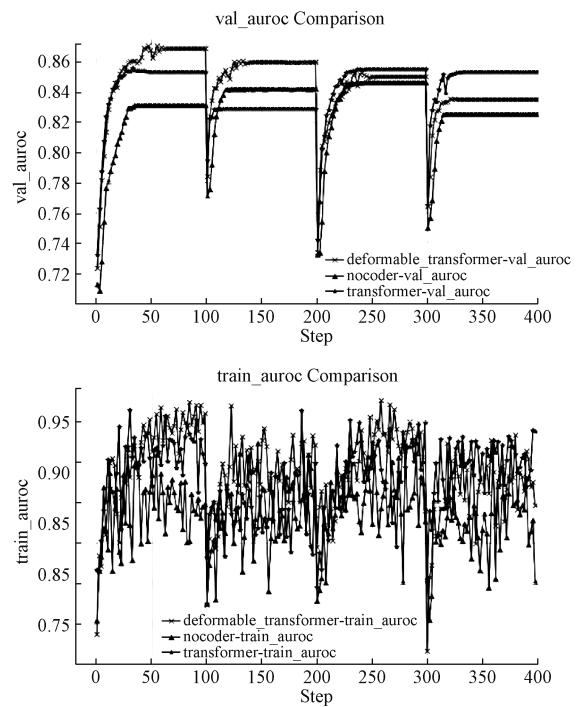


图 7 不同优化器训练集与验证集的 AUPRC 值
Fig. 7 AUPRC values of training set and validation set for different optimizers

4.3 优化器对比实验

为了体现出更好训练效果,将 Adam 优化器与本文的

AdamW 优化器进行对比实验。使用 P19 数据集,消融实验指标如表 3 所示。不添加数据噪声,设置 epoch 为 50 且 Batchsize 为 128。图 8 所示为训练集与验证集的 Loss 值,可知,无论是在训练集还是在验证集上,AdamW 优化器都明显优于 Adam 优化器,而在验证集上,损失 AdamW 比 Adam 降低了 1.8%。图 9 所示为训练集与验证集的 AUROC 值对比,可以看出,AdamW 的 AUROC 值比 Adam 高出 1.5%。图 10 所示为训练集与验证集的 AUPRC 值对比,可以得知,AdamW 的 AUPRC 值比 Adam 高出 0.4%。

表 3 P19 数据集优化器消融实验数据指标
Table 3 P19 data set optimization ablation
experimental data indicators

模型	Loss	AUROC/%	AUPRC/%
AdamW	0.559	88.2±1.8	53.7±3.6
Adam	0.562 8	87.8±1.8	52.9±3.5

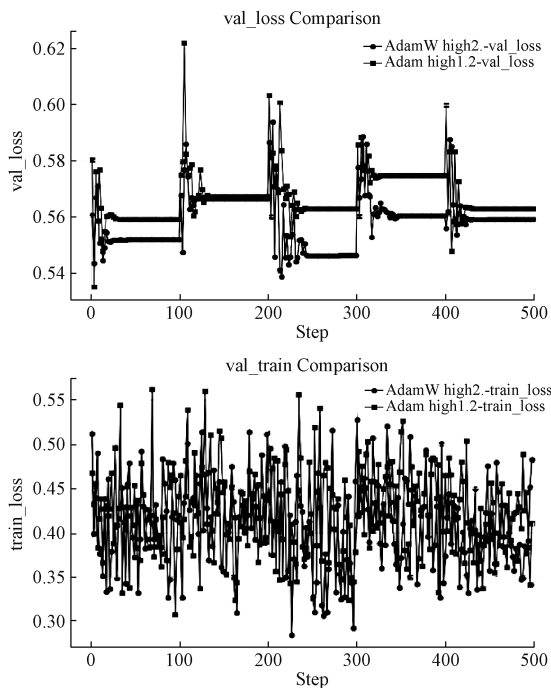


图 8 不同优化器训练集与验证集的 Loss 值
Fig. 8 Loss values of training set and validation set
of different optimizers

GDTN 模型在经过 50 个 epoch 之后达到了收敛状态,表现良好,在 P12、P19 数据集的二分类任务中,GDTN 的 AUROC 和 AUPRC 指标分别提高了 2.2%、2.6%,在 PAM 的多分类数据集中,Accuracy、Precision、Recall、F1 等指标相较于 Baseline 分别提高了 2.1%、1.1%、2.1%、2.1%。GDTN 模型表现出良好的记忆能力,能够学习到时间序列数据中的长期依赖关系,从而避免了梯度消失和

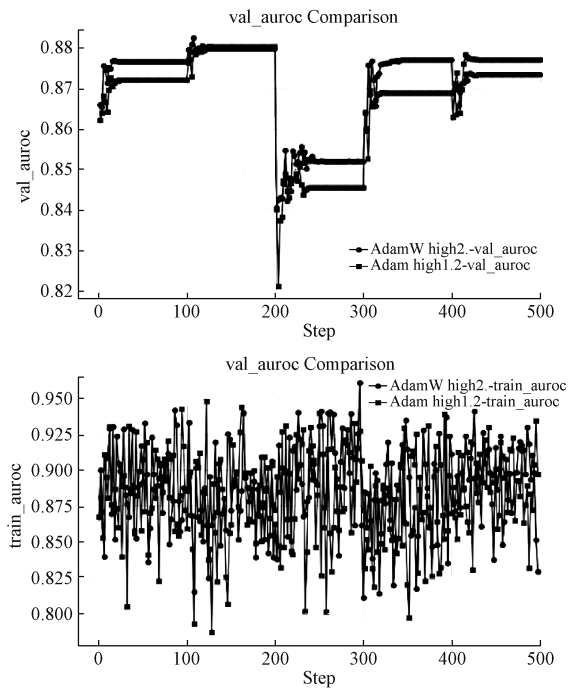


图 9 不同优化器训练集与验证集的 AUROC 值
Fig. 9 AUROC values of different optimizer training sets
and validation sets

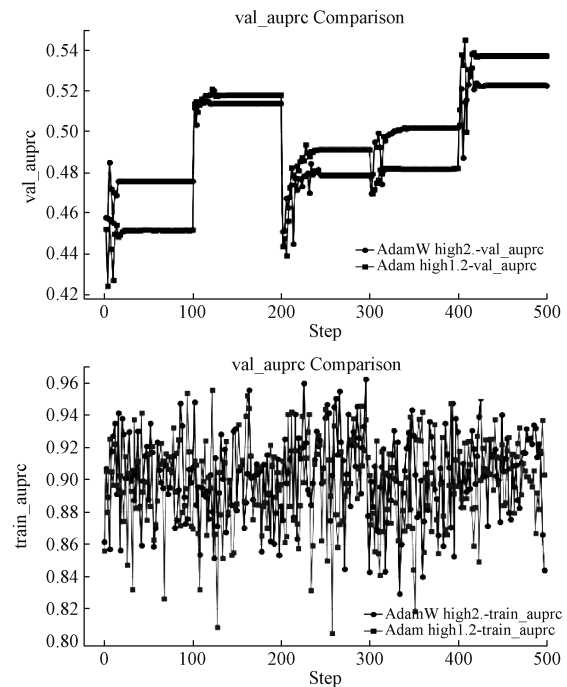


图 10 不同优化器训练集与验证集的 AUPRC 值
Fig. 10 AUPRC values of different optimizer training sets
and validation sets

梯度爆炸等问题,此外,GDTN 在较少计算资源和时间的情况下就能够达到收敛状态,进一步证明了其泛化能力

强,符合实际检测需求。

由表1可知,相对于前述的6种网络,GDTN在生命体征监测任务中表现出更优异的监测性能。而表2和3的消融实验结果则进一步说明了GDTN的算法优势,通过图5~10的可视化结果,可以清晰地观察到GDTN的性能提升。

5 结论

在生命体征监测研究任务中,Transformer、Transmean、GRU-D、SeFT、mTAND、RainDrop等处理算法致力于对于特征的提取,这6类模型主要分为Transformer类型、RNN递归类型以及常微分类模型,这些方法过程繁琐,忽略了规则时间序列对算法性能的影响,本文提出一种GDTN模型,该模型通过对人体信息采样,实现了对病人的生命体征数据的有效预测,本文将其与前述6个模型进行了详细对比。

针对时间序列的GNN算法在医疗检测中在固定时间内对患者病情预测准确度不高等问题,本文提出了一种新的预测模型——GDTN。该模型在3个不同数据集中表现均超越其他模型,展现出较强的鲁棒性,GDTN利用GNN的优势,使得网络模型能够有效处理采样到的不同时间长度信息。在医疗领域,GDTN为病人的生命体征监测提供了一种新颖、有效的方法。

在实际应用中,本文的模型仍存在一定局限性。使用的生物医学数据的尺度均为规则数据,将其应用于不同的生物医学数据尺度下的实验可能导致体征数据整合能力不佳,准确度下降问题。针对此问题,未来的研究将探索更为有效的体征数据对齐方式。考虑在数据预处理阶段使用归一化操作,规范不规则的体征数据,可利用稀疏注意力方式提高监测速度。

参考文献

- [1] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: Continual prediction with LSTM[J]. *Neural Computation*, 2000, 12(10), 2451-2471.
- [2] TAN Q, YE M, YANG B, et al. DATA-GRU: Dual-attention time-aware gated recurrent unit for irregular multivariate time series[C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(1): 930-937.
- [3] ZHANG X, ZEMAN M, TSILIGKARIDIS T, et al. Graph-guided network for irregularly sampled multivariate time series[C]. *International Conference on Learning Representations*, 2021.
- [4] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. *Advances in Neural Information Processing Systems*, 2017, 12(4): 5998-6008.
- [5] SCHAFER J L, GRAHAM J W. Missing data: Our view of the state of the art [J]. *Psychological Methods*, 2002, 7(2): 147-177.
- [6] FRITZ B A, CUI Z, ZHANG M, et al. Deep-learning model for predicting 30-day postoperative mortality[J]. *British Journal of Anaesthesia*, 2019, 123(5): 688-695.
- [7] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable transformers for end-to-end object detection[C]. *International Conference on Learning Representations*, 2020.
- [8] HE K, ZHANG X, REN S Q, et al. Deep residual learning for image recognition [C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015: 770-778.
- [9] REYNA M A, JOSEF C, SEYEDI S, et al. Early prediction of sepsis from clinical data: The PhysioNet/Computing in cardiology challenge 2019 [C]. *2019 Computing in Cardiology (CinC)*. IEEE, 2019: 1-4.
- [10] GOLDBERGER A L, AMARAL L A N, GLASS L, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals[J]. *Circulation*, 2000, 101(23): 215-220.
- [11] REISS A, STRICKER D. Introducing a new benchmarked dataset for activity monitoring [C]. *2012 16th International Symposium on Wearable Computers*, 2012.
- [12] ZHANG N, MA S, LI X, et al. Multi-scale neighborhood attention transformer on U-Net for medical image segmentation [C]. *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2022: 1381-1386.
- [13] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale [C]. *International Conference on Learning Representations*, 2020.
- [14] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021: 10012-10022.
- [15] LOSHCHELOV I, HUTTER F. Decoupled weight decay regularization[C]. *International Conference on Learning Representations*, 2017.
- [16] CHE Z, PURUSHOTHAM S, CHO K, et al. Recurrent neural networks for multivariate time series

- with missing values[J]. Scientific Reports, 2018, 8(1): 6085.
- [17] HORN M, MOOR M, BOCK C, et al. Set functions for time series [C]. International Conference on Machine Learning. PMLR, 2020: 4353-4363.
- [18] SHUKLA S N, MARLIN B. Multi-time attention networks for irregularly sampled time series [C]. International Conference on Learning Representations, 2021.
- [19] 李明, 赵文仓, 秦文谦. 基于图神经网络的任务驱动元学习方法[J]. 电子测量技术, 2021, 44(16): 123-129.
- LI M, ZHAO W C, QIN W Q. Task-driven meta-

learning method based on graph neural network [J]. Electronic Measurement Technology, 2021, 44(16): 123-129.

作者简介

孙佳琪, 本科, 主要研究方向为人工智能、图像处理。
E-mail: 1623958071@qq.com

单慧琳(通信作者), 博士, 副教授, 主要研究方向为深度学习、图像处理等。
E-mail: shanhuilin@nuist.edu.cn