

基于麻雀搜索算法的混合神经网络模型及其血糖预测应用

徐 鹤^{1,2}, 许硕洋¹, 季一木^{1,2}

(1. 南京邮电大学计算机学院、软件学院、网络空间安全学院, 南京 210023; 2. 江苏省高性能计算与智能处理工程研究中心, 南京 210023)

摘 要: 糖尿病是当今危害人类健康的常见疾病之一, 有效管理和控制血糖对患者至关重要。传统的血糖预测模型大多为单一的深度学习模型, 存在精度不足或效率太低的缺陷, 制约了其在实际应用中的效果, 为此, 本文提出了一种基于麻雀搜索的混合神经网络模型, 将其应用到血糖预测中。该模型结合了时域卷积网络(Temporal convolutional network, TCN)和门控循环单元(Gated recurrent unit, GRU), 是基于端到端方式训练的时序神经网络, 根据患者的血糖水平历史记录预测血糖。为确保该模型的泛化能力, 使用两个不同来源的数据集进行验证。首先, 对多源时序监测数据的特征采样频率进行设定, 时间间隔为 5 min, 接着对数据做平滑处理和标准化, 并通过 TCN 对时序数据捕捉时序模式和依赖特征; 然后通过构建基于注意力机制的 GRU (GRU-Attention) 模型进一步提取特征并建模; 最后使用麻雀搜索算法对 TCN 和 GRU-Attention 模型进行超参数优化, 实现血糖预测模型。为了证明本文所提模型的有效性, 将其预测结果与其他模型进行对比, 包括 LSTM、ARIMA 和 RNN 等。研究结果表明, 提出的基于麻雀搜索算法的 TCN 和 GRU-Attention 模型在血糖值预测任务中表现良好, 两个数据集的均方根误差(Root mean square error, RMSE)和平均绝对误差(Mean absolute error, MAE)分别为 0.552 和 0.402、0.531 和 0.388, 均优于其他模型。

关键词: 血糖预测; 麻雀搜索算法; 注意力机制; 时域卷积网络; 门控循环单元

中图分类号: TP391

文献标志码: A

Hybrid Neural Network Model Based on Sparrow Search Algorithm and Its Application in Blood Glucose Prediction

XU He^{1,2}, XU Shuoyang¹, JI Yimu^{1,2}

(1. School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2. Jiangsu HPC and Intelligent Processing Engineer Research Center, Nanjing 210023, China)

Abstract: Diabetes is one of the most common diseases that endanger human health. Effective management and control of blood sugar is very important for patients. Traditional blood glucose prediction models are mostly single deep learning models, which have the defect of insufficient accuracy or low efficiency, restricting their effect in practical application. Therefore, a hybrid neural network model based on sparrow search is proposed and applied to blood glucose prediction. The proposed model combines a temporal convolutional network (TCN) and a gated recurrent unit (GRU), and it is a sequential neural

基金项目: 江苏省科技支撑计划项目(BE2019740); 江苏省六大人才高峰高层次人才项目(RJFW-111); 江苏鱼跃医疗设备股份有限公司科技项目(2022外017); 江苏省研究生实践创新计划项目(SJCX23_0274)。

收稿日期: 2024-01-15; **修订日期:** 2024-04-11

network trained in an end-to-end manner to predict blood glucose based on a patient's blood glucose level history. In order to ensure the generalization ability of the model, data sets from two different sources are used for validation. Firstly, the feature sampling frequency of multi-source timing monitoring data is set at a time interval of 5 min, the data is smooth-processed and standardized, and the timing pattern and dependency characteristics are captured by TCN. Then, by constructing a GRU model based on the attention mechanism (GRU-Attention), features are further extracted and modeled. Finally, the sparrow search algorithm is used to optimize the hyperparameters of the TCN and GRU-attention models to realize the blood glucose prediction model. To prove the validity of the proposed model, its prediction results are compared with those of other models, including LSTM, ARIMA, RNN, etc. The results show that the proposed TCN and GRU-Attention models based on the sparrow search algorithm perform well in the task of predicting blood glucose value. The root mean square error (RMSE) and mean absolute error (MAE) of the two datasets are 0.552 and 0.402, 0.531 and 0.388, respectively, which are all better than other models.

Key words: blood glucose prediction; sparrow search algorithm; attention mechanism; temporal convolutional network (TCN); gated recurrent unit (GRU)

引言

糖尿病作为一种慢性的代谢性疾病,主要源于人体细胞膜上胰岛素受体数量减少或受体敏感性下降^[1],导致胰岛素的分泌无法满足生理需要。由于胰岛素的绝对或相对不足,使其无法有效发挥生理作用,进而导致患者血糖难以维持在正常范围^[2]。这一病理对全球流行病学和公共卫生构成了严峻挑战。在这一背景下,确保精准的血糖水平预测对于糖尿病患者的生活质量至关重要。精准的血糖预测有助于提前干预和调整治疗方案,以有效控制血糖水平。因此,发展和应用可靠的血糖预测模型成为科学领域中研究的关键方向之一。

为了提升血糖预测的精度,国际国内众多专家学者一直致力于研究高效的预测模型,使用最广泛的包括机器学习方法和回归分析方法^[3]。近年来,随着人工智能领域算法技术的不断迭代和深度学习的崛起,一系列基于神经网络的方法也正逐渐成为糖尿病血糖预测领域的焦点。这些深度学习模型能够更灵活地处理复杂的时空动态特征^[4]和非线性关系。以基于线性关系假设的差分自回归滑动平均(Autoregressive integrated moving average, ARIMA)模型为例,对于某些复杂的时间序列难以捕捉到非线性动态特征^[5],且对时间序列的平稳性要求^[6]较高。相比之下,长短期记忆网络(Long short-term memory, LSTM)尽管在序列数据处理方面表现出色,但由于其对内存的大量需求,需要存储隐藏状态,且后续时间的预测受限于前一次的完成,因而训练过程比较耗时。

深度学习方法的引入虽然为糖尿病血糖预测提供了新思路,但也需要权衡计算资源消耗与模型性能,特别是在要求快速响应的医疗应用^[7]中。同时,不难发现,在当前的血糖预测研究实验中,大多数学者都倾向采用单一的神经网络模型。然而,这种做法在捕捉时间序列信息方面存在诸多限制^[8],对于序列数据中复杂时空动态关系的挖掘能力不足^[9],难以有效处理时间序列数据中的动态变化,从而引发模型精度不足的问题。

为应对上述问题,需要考虑采用更为复杂的模型结构,例如结合多种神经网络架构。对此,本研究提出了一种基于麻雀搜索算法的时域卷积网络(Temporal convolutional network, TCN)和门控循环单元(Gated recurrent unit, GRU)模型,并将其应用于血糖预测中。该模型通过麻雀搜索算法对TCN和

GRU模型进行超参数调优,优化模型的性能。在训练过程中,TCN并行计算的特性^[10]可以加速模型训练和推断的过程,用于捕捉数据的时序特征;在基于非线性拟合的GRU中引入注意力机制(Attention mechanism),有助于GRU在快速建模的过程中更好地聚焦时序关系^[11],提升模型对关键信息的捕获。基于此,本研究基于机器学习理论在血糖预测领域的贡献如下:

(1) 通过采用麻雀搜索的群智能优化算法,利用群体智能的优势,通过多个个体之间的信息共享和协同搜索,对TCN和GRU模型进行超参数优化,提高搜索效率和全局优化的能力,以实现最优性能。

(2) 结合TCN和GRU两种不同深度学习模型,充分利用TCN的并行计算特性,学习过去的时序数据中的依赖关系,得到原始数据的时序特征,然后在GRU中增加Attention层,在数据建模过程中进一步挖掘时序中的关键信息,应用到血糖预测任务中,提高血糖预测效率和准确性。

(3) 在数据预处理阶段,通过滑动窗口的方式对时间序列数据做平滑处理,减少数据波动,更有效地捕捉时间序列中的趋势和周期性变化,提高模型对于数据特征的敏感性。

(4) 对两种数据集的多名不同患者的血糖数据进行神经网络模型的学习和训练,以实现对患者血糖水平的准确预测。

1 相关工作

在实际临床应用中,对血糖预测的准确性要求很高。研究人员在开发血糖预测模型领域研究了大量工作,随着深度学习技术的不断成熟,神经网络在血糖值和糖尿病预测任务中应用越来越受欢迎。以LSTM为例,Rabby等^[12]使用基于LSTM的循环神经网络模型进行血糖预测的深度学习,在优化最佳特征组合的基础上,通过堆叠多层LSTM提高血糖预测精度;Yang等^[13]提出了一种自回归积分移动平均模型,基于模型自适应识别算法进行血糖预测和低血糖预警;Martinsson等^[14]讨论使用循环神经网络(Recurrent neural network, RNN)模型,通过学习患者的血糖水平历史记录来预测未来的血糖水平,并在建模过程中使用LSTM单元优化以应对针对RNN的梯度消失问题;Peng等^[15]针对LSTM参数太多、容易过拟合的缺陷,在简化LSTM网络规模的基础上融入GRU,构建LSTM-GRU模型预测血糖水平,性能表现良好。

针对上述现有的血糖预测方法,无论是改进的单一预测模型还是混合预测模型,都只擅长在短期内获得良好的预测效果。然而,血糖水平的变化通常具有强烈的时序依赖性,当前或未来时刻的血糖水平受过去时刻的影响很大。在实际临床应用中,特别是长时间范围内的时序血糖数据,医疗专业人士和患者可能很难理解模型的预测结果。为了获得更佳的预测效果,很多研究人员使用注意力机制增强长期依赖关系的捕获效果。

注意力在增加神经元之间的连接方面具有很好的特点,因为它可以自适应地选择相关的隐藏状态并在每个时间步维持所有隐藏状态。在传统的神经网络中,每个神经元的输出只依赖于前一层的所有神经元的输出,而在注意力机制中,每个神经元的输出不仅仅取决于前一层的所有神经元的输出^[16],还依赖于数据不同部分的权重,能够让模型更加关注输入序列中的关键信息,从而提高模型的精度和效率。

人类视觉在信息接收的过程中并非是一蹴而就的,而是通过将注意力聚焦在特定部分上,以“高分辨率”感知关注区域,而对其他部分进行“低分辨率”感知^[17]。这意味着通过赋予特定部分更高的关注权重,以极小的关注权重感知其他部分。注意力机制被设计成模仿人类大脑和视觉系统处理信息的方式,通过为不同信息赋予不同的权重。Yan等^[18]提出一种基于时空注意力机制的编码器-解码器神经网络结构,考虑视频中的空间和时间结构,使解码器能够自动选择最相关的时间段中的重要区域进行单

词预测;Ran等^[19]使用LSTM对复杂非线性时空交通动态进行建模,将注意力机制应用于每个LSTM单元的输出层,所提出的模型比单一的LSTM模型具有更好的准确性;Zhang等^[20]构建了基于深度分解机和注意力机制的卷积神经网络模型,通过增强的特征学习来提高股票价格走势的预测精度;Chen等^[21]针对机械中的关键部分零件的寿命预测,构建了一种基于注意力机制的轴承剩余使用寿命预测的深度学习新方法,取得了很好的性能。

但注意力机制并没有被广泛地应用在实际的血糖预测任务中。本文将结合GRU和注意力机制构建深度学习神经网络模型GRU-Attention,并应用到血糖预测应用中。注意力机制会在GRU捕捉时序特征的基础上更加聚焦于具有更高权重的特定信息,同时抑制对于其他无关信息的过多关注,使得GRU在处理序列数据^[22]时更为准确和高效,从而提高血糖预测任务的精确性。

同时,构建有效的深度学习模型是一个复杂且耗时的过程,这涉及深度学习模型的适当结构及其超参数的不断调整寻优,并且模型的准确性很大程度上依赖于超参数的选择。在大数据时代之前,神经网络模型的结构设计相对简单,模型的超参数通常可以手动调整。然而,随着深度学习技术的不断进步,各种神经网络模型变得越来越复杂,其超参数的数量也不断增加,手动调整参数变得愈发繁琐且不切实际。为了解决这个问题,研究专家们提出了群体智能优化算法^[23]进行超参数优化。

群体智能算法能在每次迭代中学习,并在下一次迭代中向更好的超参数值迈进。一些著名的群体智能算法包括粒子群优化和麻雀搜索优化等^[24]。Wang等^[25]提出了一种结合变分模式分解和改进的粒子群优化长短期记忆网络的短期血糖预测模型,该模型仍具有较高的预测精度;An等^[26]讨论了使用深度极限学习机进行风电功率预测,为解决模型输入权重和阈值随机变化的问题,引入麻雀搜索算法优化其参数,实验结果有较好的提升;Bui等^[27]为了验证群体智能算法对深度学习神经网络优化的有效性,实验比较了多种模型进行洪水敏感性绘图,结果表明,所提出的群体智能模型在几乎所有比较指标上都优于基准方法。

麻雀搜索算法(Sparrow search algorithm, SSA)是一种基于自然麻雀群体行为的群体智能优化算法。该算法是通过模拟麻雀种群的寻找食物、社交和逃避捕食行为,找到最优位置^[28-31]。本文选择基于群智能优化的麻雀搜索算法作为TCN和GRU-Attention混合神经网络模型的超参数寻优方法,并将其应用到血糖预测任务中。

2 模型架构的提出

本研究构建了一种基于麻雀搜索算法的混合神经网络模型,为端到端的时序神经网络模型提供了一种新的训练方法。它的核心理念在于运用TCN捕获时间序列数据在不同尺度下的关键特征,然后使用GRU-Attention深度学习模型对数据变量进行非线性拟合^[32]。通过引入麻雀搜索算法,对TCN和GRU-Attention模型进行超参数优化,最终实现血糖值预测。

图1展示了本文提出的基于麻雀搜索算法的混合神经网络模型的整体流程和结构,主要包括数据的预处理、输入层、神经网络模型的构建以及输出层,其中模型构建分为3个关键部分:

(1) 利用TCN捕获时间序列数据组成的输入张量在不同尺度下的关键特征,然后将TCN对时间序列特征的高效学习用于后续GRU的建模。

(2) 在获取到的时序特征的基础上,通过GRU-Attention模型中处理时序数据的长期相关性,利用GRU的快速建模能力进行非线性拟合,预测血糖值。

(3) 引入麻雀搜索算法,对TCN和GRU-Attention进行参数优化,进一步提升血糖预测的精度。

本文所提出的模型名为基于SSA的TCN-GRU-Attention,简称为S-T-GA。

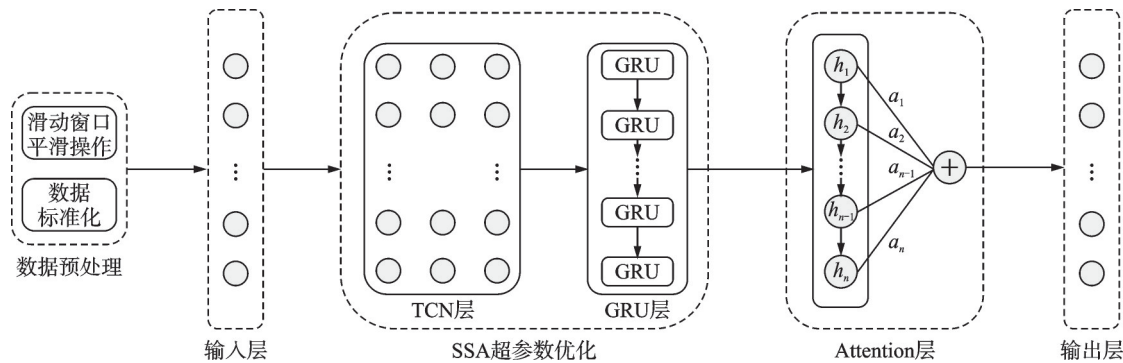


图1 基于麻雀搜索算法的混合神经网络模型结构图

Fig.1 Structural diagram of hybrid neural network model based on sparrow search algorithm

2.1 数据预处理

2.1.1 多源时序监测数据融合与频率调整

由于本研究使用的公开血糖数据涉及两个设备的监测和采集,而每个设备采集到的特征具有不同的采样频率,为了统一时间维度以便进行有效的分析,首要任务是对所有特征的时间序列进行调整,使其拥有相同的采样时间点。首先,将两个设备采集到的数据进行合并。考虑到数据融合过程中每个特征在时间上的不对齐,使用降采样技术将多元数据进行时序对齐。这意味着在整合数据时,为确保所有特征匹配到相同时间点,通过将数据在时间轴上进行降采样的方式,使得所有数据特征具有一致的时间频率。最终,将所有特征的时间序列调整为每5 min进行一次采样,以保证数据的时序一致性。

该数据融合与频率调整的过程通过有效整合来自不同设备的监测数据,为后续的深度学习模型分析提供了一个统一可靠的时间基准,确保了多源时序数据的一致性和可比性。

2.1.2 数据平滑处理

在实际生活和糖尿病诊治过程中,血糖值通常受到多种因素的影响,包括心率、胰岛素、膳食摄入量、身体活动以及不同时间等动态因素,并且在采集血糖数据的过程中,血糖监测仪或其他可穿戴设备可能会因为患者不同的生活活动情况而导致非线性的血糖数据出现很大的波动。如果将不平稳的数据直接输入深度学习模型中,神经网络很难学习到数据的长期模式。因此,本文首先对数据做平滑处理操作。常见的数据平滑处理方法有滑动平均、多项式拟合、Kalman滤波和多项式拟合等。在医学监测场景中,简单直观的滑动平均法相较于其他复杂的平滑处理在实时性方面更具有优势。

滑动平均法是一种在时间序列分析任务中常见的数据平滑技术^[33],其基本思想是通过计算连续时间窗口内数据的平均值,不断右移加入新数据点,在每次滑动后再重新计算窗口内数据的平均值,以此来平滑时序数据。可以通过调整滑动窗口的大小权衡平滑效果和灵敏度,减少时序数据中的短期波动和噪声,从而反映出数据的整体趋势。窗口内数据平滑处理的计算公式为

$$y_t = \frac{x_{t-k-1} + x_{t-k-2} + \cdots + x_t}{k} \tag{1}$$

式中: x_t 表示在 t 时刻的原始数据值, k 为滑动窗口的大小。

2.1.3 数据标准化

因为不同特征具有不同的度量单位和数值范围,这种情况可能导致不同特征对模型的影响程度各异。为了消除这些特征之间的尺度差异,可以通过标准化操作使得所有特征都在相似的尺度上,提高模型的稳定性和收敛性。常用的标准化方法有鲁棒标准化、单位向量标准化、Min-Max标准化、对数变换和标准差标准化。

在本研究中,使用Z-score标准化,这是一种线性变换^[34]的方法,通过对每个特征进行平均值的中心化,再除以标准差,处理后的特征具有均值为0、标准差为1的特性,且不会改变原始数据的分布形状,保持了数据的统计特性^[35],其计算公式为

$$Z = \frac{X - \mu}{\sigma} \quad (2)$$

式中: Z 表示标准化处理后的数据, X 表示原始数据, μ 表示数据的均值, σ 表示数据的标准差。

2.2 T-GA 回归预测模型

如图1所示,T-GA回归预测模型主要包括5部分:输入层、TCN层、GRU层、Attention层和输出层。

(1) 输入层

根据数据预处理中所述的方法,对数据做融合与频率调整、平滑和标准化操作,将预处理后的血糖数据作为TCN的输入。

(2) TCN层

TCN在卷积神经网络^[36]的基础上克服了不能处理时序问题的短板,由扩张因果卷积核(Dilated causal convolutional, DCC)和残差连接部分(Residual connections, RC)组成。

将预处理完成之后的血糖数据输入TCN中,第一层卷积会对输入序列做时序特征提取,且随着逐层扩张来捕获不同尺度下的血糖动态模式,这种机制也确保了模型能够摒弃对未来数据的信息解读,避免数据泄露。然而,随着层数的不断增多,梯度消失或梯度爆炸的问题也随之而来,对此,TCN通过残差结构让相邻的两层卷积之间发送误差信息,促进信息流动并加快模型的收敛速度。

TCN的扩张因果卷积为单向传输时序数据,只对当前和以前的血糖时间步长执行卷积运算,计算公式为

$$y_t = \sum_{i=0}^k x_{t-i} * w_i \quad (3)$$

式中: y_t 表示时间 t 的输出, x 表示输入序列, $t-i$ 表示与卷积核中元素对应的序列, w 表示卷积核, k 表示卷积核的大小。

因为血糖水平的波动可能在较长时间尺度上产生相关性,因此可以在计算复杂度不变的前提下,通过增加卷积核大小和扩张系数,接受更广泛的血糖输入信息,增大感受野,计算公式为

$$y_t = \sum_{i=0}^k x_{t-d \times i} * w_i \quad (4)$$

式中: d 表示扩张系数,决定卷积核中每个权重之间的间距。如图2(a)所示扩张因果卷积结构, $d=1$ 时表示在输入层对每个点进行采样, $d=2$ 时表示每2个点作为输入进行采样, $d=4$ 时表示每4个点作为输入进行采样。扩张系数会随着层的增加呈指数级增加,因此网络的更深层可以捕获到更大范围的时序关系。

在TCN的残差结构中,如图2(b)所示残差块结构,每一个基本残差模块都由2个卷积层和1个直接映射部分组成。为规范网络,每一层中也都有WeightNorm、ReLU激活函数和Dropout。同时为了保证输入和输出的形状一致,通过额外的一维卷积确保求和运算中的两个张量形状一致,以得到最终输出。

(3) GRU层

基于TCN训练学习血糖数据得到的时序特征,使用GRU提取时序依赖关系。GRU是循环神经网络(Recurrent neural network, RNN)的一个重要变体,在时间序列上具有很强的长记忆能力,与LSTM相比,其网络结构大大简化,训练率显著提高。

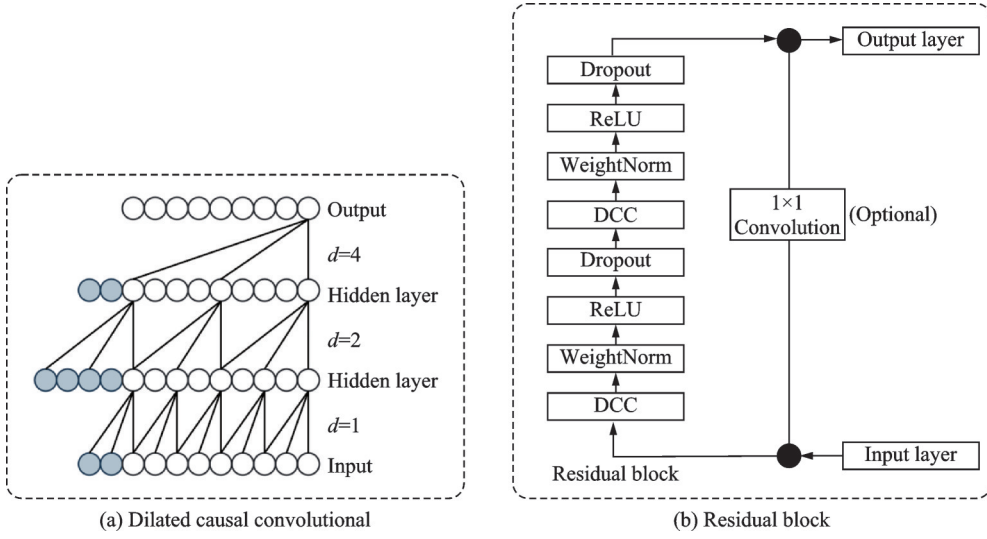


图2 TCN模型结构

Fig.2 Structure of TCN model

如图3所示,GRU主要结构^[37]为更新门和重置门。更新门控制前一时间步隐藏状态对当前时刻隐藏状态的影响程度,将上一时间步的隐藏状态 h_{t-1} 和当前时间点的输入 x_t 拼接后做线性变换,再通过 Sigmoid 激活函数计算出更新门控信号;重置门决定上一时刻隐藏状态信息的保留程度^[38],计算方式与更新门相似,但结果会作用在上一时刻的隐藏状态,并通过 Tanh 激活函数计算出新的隐藏状态。最后将更新门和重置门的输出结果相加,得到最终的隐藏状态。其中,计算候选隐藏状态和当前时间步隐藏状态可分别表示为

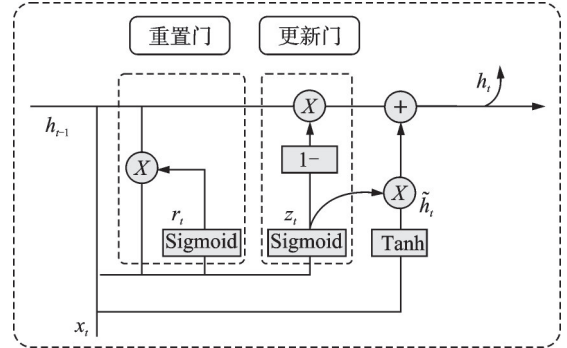


图3 GRU结构

Fig.3 GRU structure

$$z_t = \text{Sigmoid}(W_{xz} \cdot x_t + W_{hz} \cdot h_{t-1} + b_z) \quad (5)$$

$$r_t = \text{Sigmoid}(W_{xr} \cdot x_t + W_{hr} \cdot h_{t-1} + b_r) \quad (6)$$

$$\tilde{h}_t = \text{Tanh}(W_{xh} \cdot x_t + W_{hh} (r_t \cdot h_{t-1}) + b_h) \quad (7)$$

$$h_t = (1 - z_t) \cdot \tilde{h}_t + z_t \cdot h_{t-1} \quad (8)$$

式中: z_t 和 r_t 分别表示为更新门和重置门; Sigmoid 表示激活函数; x_t 表示输入向量; W 表示权重矩阵; b 表示偏差; Tanh 表示双曲切线激活函数; $\tilde{h}_t$ 和 h_t 分别表示时间步 t 的候选隐藏状态和隐藏状态。

(4) Attention层

随着时间步长的不断增加,GRU在捕获长距离时序依赖关系方面可能面临挑战。为了增强模型对时序血糖数据中关键信息的动态捕捉能力,引入注意力机制以灵活地关注输入序列中的不同时间点,这在实际医学诊治过程中是可靠的。

在计算隐藏状态时,根据输入序列的不同时间步动态分配权重,赋予重要时间步更高的关注度,从而提升GRU模型对全局相关性信息的有效利用率。这种动态的权重分配机制使得模型能够更灵活地

适应不同时间步之间的关系,从而更有效地捕获长期时间序列中的重要模式和依赖性。

将GRU层计算出的 t 时刻的隐藏状态 h_t 作为输入,计算公式^[39]表示如下

$$e_t = \text{Tanh}(Wh_t + b) \quad (9)$$

$$a_t = \frac{\exp(e_t)}{\sum_{j=1}^T \exp(e_j)} \quad (10)$$

$$H_t = \sum_{i=1}^T a_i h_i \quad (11)$$

式中: e_t 表示注意力激励值; T 表示时间步数; $\exp(\cdot)$ 为指数函数; a_t 表示 e_t 的归一化后的注意力权重; H_t 表示每个时间步的隐藏状态与对应注意力权重相乘的求和。

(5) 输出层

激活全连接层的函数选取Sigmoid函数,得到 t 时刻的最终血糖预测值 Y_t 表示为

$$Y_t = \text{Sigmoid}(WH_t + b) \quad (12)$$

2.3 SSA优化T-GA回归器

实际临床应用要求高精度的血糖值预测,前文提到的TCN和GRU神经网络中多个超参数的设定对最终的血糖值预测结果影响很大。为了提高预测效果,应用麻雀搜索算法分别对TCN和GRU进行优化参数选择,适应度函数选取均方根误差(Mean absolute error, MAE)。

SSA的实现流程^[40]为:

- (1) 随机生成一组解(一群麻雀),每一个参数组合就相当于一只麻雀,初始化搜索空间,确定每个参数的取值范围;
- (2) 对每个参数组合计算出适应度函数,适应度越好,表示该参数组合效果越优秀,反映模型在血糖值预测任务中的性能;
- (3) 对于适应度比较好的参数组合会通过局部搜索细化参数组合,找到更优秀的值;
- (4) TCN和GRU模型中都设定了早停机制,步长参数为3,当达到迭代停止或适应度达到一定阈值时,停止搜索,最后输出最优适应度结果以及对应的超参数组合。

表1和表2分别展示了TCN和GRU模型的主要参数描述,这些参数是通过SSA进行优化的关键配置。该算法在搜索空间中动态调整这些参数,以最大程度地提升混合神经网络模型在血糖预测任务中的性能。

SSA优化T-GA混合神经网络模型的核心算法如下:

输入

种群规模 N ,种群位置 pos ,迭代次数 T ,发现者数量 PD ,危险感知者数量 SD ,预警值 R_2 ,安全阈值 ST 。

输出

最优位置 pos_TCN 和 pos_GRU 、最优适应度。

(1) 初始化:

(2) pos_TCN : 'num_channels': [16, 32, 64], 'kernel_size': [2, 3, 4]

(3) pos_GRU : 'hidden_size': [32, 64,

表1 TCN主要超参数描述

Table 1 Description of TCN main hyperparameters

参数	描述	范围
num_channels	通道数	[16, 32, 64]
kernel_size	卷积核大小	[2, 3, 4]

表2 GRU主要超参数说明

Table 2 Description of GRU main hyperparameters

参数	描述	范围
hidden_size	隐藏层单元数	[32, 64, 128]
num_layers	GRU层数	[1, 2, 3]
seq_len	时间步长	[10, 20, 30]


```

128]、'num_layers':[1,2,3]、'seq_len':[10,20,30]
(4) While  $t < T$ 
(5) 计算麻雀种群适应度并排序,找出当前最优和最差适应度的麻雀位置
(6)  $R = \text{rand}(1)$ 
(7) If  $R_2 < ST$ 
(8)   for  $i = 1:PD$ 
(9)     更新发现者位置 pos_TCN [ $i$ ] 和 pos_GRU [ $i$ ]
(10)    Attention (new_hidden_state) # 更新GRU隐藏状态计算 Attention输出
(11)   end for
(12)   for  $i = (PD+1):N$ 
(13)     更新跟随者位置 pos_TCN [ $i$ ] 和 pos_GRU [ $i$ ]
(14)    Attention (new_hidden_state) # 更新GRU隐藏状态计算 Attention输出
(15)   end for
(16)   for  $j = 1:SD$ 
(17)     更新危险感知者位置 pos_TCN [ $i$ ] 和 pos_GRU [ $i$ ]
(18)    Attention (new_hidden_state) # 更新GRU隐藏状态计算 Attention输出
(19)   end for
(20) end if
(21) end While
(22) Function Attention (hidden_state):
(23) 初始化注意力权重列表 attention_weights
(24) for  $i = 1:\text{inputs}$ :
(25)  计算注意力分数,添加到 attention_weights
(26)  归一化 attention_weights
(27)  对输入序列进行加权求和
(28) end for
(29) 返回输出值

```

3 实验与分析

3.1 实验数据

为验证 S-T-GA 模型的有效性和泛化性,本研究使用两个不同来源的血糖数据集,分别为某实验室公开的血糖数据集和鱼跃集团提供的传感器血糖数据集。

公开数据集是通过连续血糖监测仪和腕戴式可穿戴设备联合采集的。研究对象包括 16 名年龄在 35 至 65 岁之间、血糖升高在正常范围内的参与者,持续进行了 8~10 天的监测。连续血糖监测仪每 5 min 测量一次葡萄糖浓度(mg/dL 转 mmol/L),而腕戴式可穿戴设备则以 64 Hz 采样光电体积描记法(Photoplethysmography, PPG)、4 Hz 采样皮肤电活动(Electrodermal activity, EDA)和皮肤温度(Temp),以及 32 Hz 采样三轴加速度测量。这些采集到的数据共包括 7 个特征,其中心率(Heart rate, HR)和血容量脉冲(Blood volumn pulse, BVP)信号每秒提供一次,而心跳间隔(Interbeat interval, IBI)则由这些信号计算得到。在实验设计中,训练集选用 16 名参与者的第 1~7 天的数据,测试选用 16 名参与者在第 8~10 天的数据,使用每组超参数测试 T-GA 混合神经网络模型在实际血糖预测中的性能指标。

鱼跃集团的传感器数据包括 6 名患者 2022 年 10 月 20 日的血糖数据和 5 名患者 2022 年 11 月 1 日的

血糖数据。主要数据项包括传感器运行时间、工作电流信号值、空白电流信号值、体表温度、实际血糖值、校准值、持续血糖监测浓度、血糖计算中涉及实际电流值、血糖浓度(mmol/L)。在实验设计中,训练集选用10月20日的数据,测试集选用11月1日的数据。使用每组超参数测试T-GA混合神经网络模型在实际血糖预测中的性能指标。表3和表4分别提供了公开数据集和鱼跃数据集中的相关特征描述。

表3 公开数据集数据特征描述

Table 3 Description of data features of public dataset

数据项	描述
Timestamp	数据采样时间
acc_x, acc_y, acc_z	三轴加速度(与主体运动和姿势相关)
BVP	血容量脉冲信号
EDA	皮肤电活性
HR	心率(每分钟心跳次数)
IBI	心跳间隔
Temp	皮肤温度
Glucose	血糖值

表4 鱼跃数据集数据特征描述

Table 4 Description of data features of YuWell dataset

数据项	描述
Time	时间戳
IW	工作电流
IB	空白电流
<i>t</i>	皮肤温度
CNO	指尖采血得到的实际血糖浓度
BG	传感器校准值
IW_net	实际工作电流
Base_ref	CGM浓度
Ref	未经校准的血糖值

图4和图5分别展示了两个数据集的散点图矩阵。散点图呈现了两两特征之间的线性关系,其中横、纵坐标分别代表数据集中的各特征;而对角线上的直方图反映了各特征变量的值分布情况,其中横坐标表示数据的取值范围,纵坐标表示对应取值范围内的样本数量或频率。

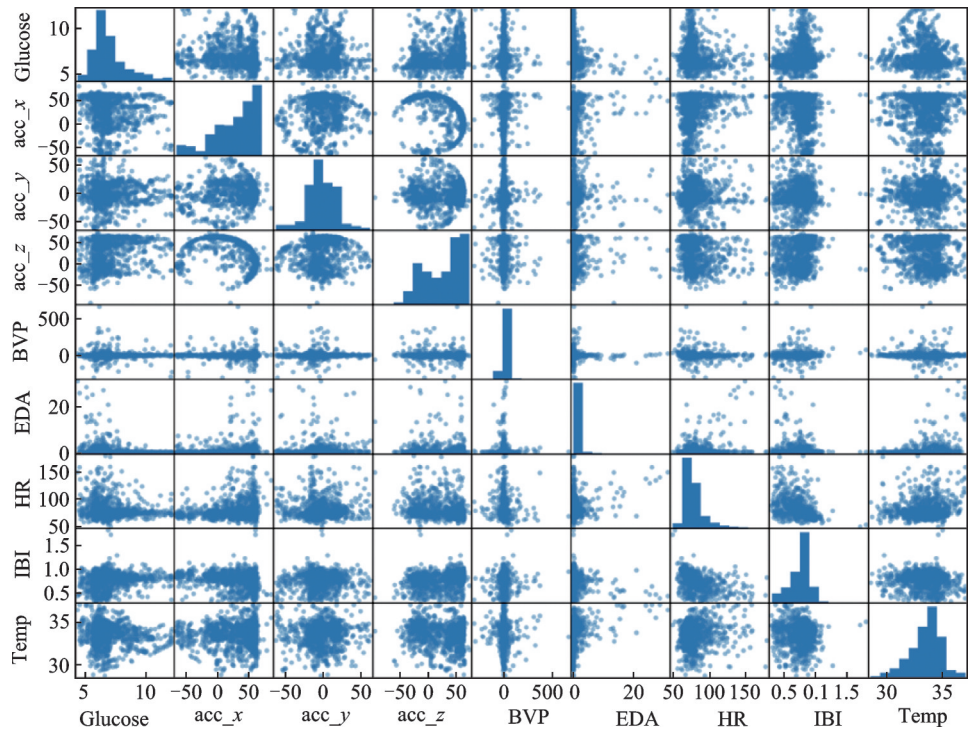


图4 公开数据集的散点图矩阵

Fig.4 Scatter plot matrix of public dataset

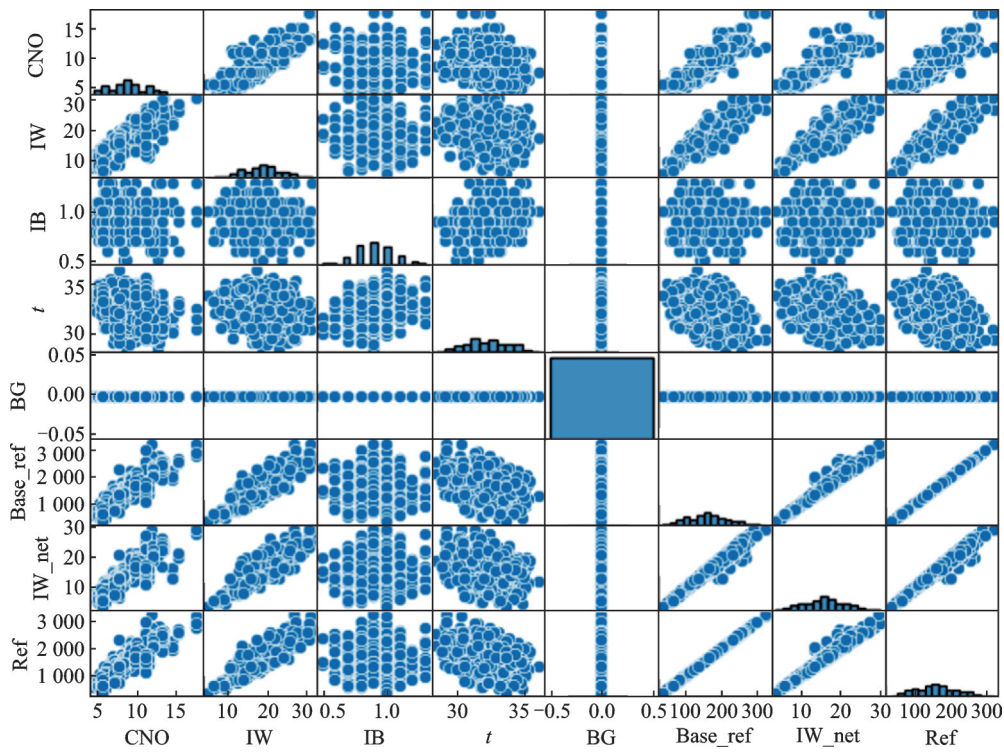


图5 鱼跃数据集的散点图矩阵

Fig.5 Scatter plot matrix of YuWell dataset

观察图4可以发现,公开数据集的特征BVP、EDA的直方图呈现较为聚集的分布,HR、IBI、Temp与其他特征的相关性散点分布不规律,有明显的非线性关系。同时,在图5中可以观察到鱼跃数据集中的特征CNO与IB和 t 呈现出很强的非线性关系。由于GRU模型有较强的非线性建模能力,因此可以捕捉到特征之间的复杂关系,更准确地拟合数据中的复杂模式。

另外,在特征相关性方面,公开数据集中的Glucose与其他特征之间普遍呈正相关性,EDA与HR、Temp等特征之间存在较强的相关性,说明生理信号之间可能存在一定的相互关联。Glucose与IBI、acc_z之间的相关性系数也相对较高。但通常情况下,三轴加速度测量用来捕捉和描述人体活动模式。因此,公开数据集选择的训练特征以生理信号为主,包括BVP、EDA、HR、Temp、IBI。数据采样频率为5 min一次,基于这些关键特征实现更精准的血糖值预测。

鱼跃数据集中的IB与CNO之间具有低相关性,IW、IW_net、Base_ref、Ref与CNO之间存在较强的相关性,但是Base_ref和Ref是由传感器内部计算的输出和校准值,因此不能作为训练特征。因此,鱼跃数据集选择的训练特征包括IW、IB、IW_net和 t 。数据频率为3 min一次,基于这些关键特征实现更精准的血糖值预测。

为了综合考虑特征之间的多维相关性和相互影响,而不仅仅是特征与目标变量之间的关系,通过TCN的扩张卷积初次提取序列数据中的特征,并以多层堆叠的方式增强模型的表示能力,能够更好地捕捉到数据中的复杂模式和关系,从而提高血糖预测的准确性和可靠性。

3.2 实验设置

本研究的实验使用的计算机配置为M2芯片,八核CPU,八核GPU,16G RAM,macos Sonoma操

作系统;使用编程语言 Python3.10,编程环境 Pycharm,使用 Pandas、NumPy、Matplotlib 等库和 PyTorch 深度学习框架,使用 GPU 加速计算。

3.3 评价指标

针对模型评估,本文使用 RMSE 和 MAE 两种评估指标测试 S-T-GA 模型的性能。RMSE 和 MAE 的计算公式分别为

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_{\text{tru}}(i) - y_{\text{pre}}(i))^2} \quad (13)$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_{\text{tru}}(i) - y_{\text{pre}}(i)| \quad (14)$$

式中: y_{tru} 和 y_{pre} 分别表示样本 i 的真实值和预测值, N 表示样本数。

3.4 实验设计与结果分析

为了验证本文提出的 T-GA 混合模型本身的优越性,通过计算多种模型的 RMSE 和 MAE,比较了不同算法在两个数据集上的预测效果,包括 TCN-GRU、TCN-LSTM、LSTM、ARIMA 和 RNN 模型。预测结果如表 5 所示。

表 5 不同深度学习模型的血糖预测实验结果对比

Table 5 Comparison of blood glucose prediction experiment results by different deep learning models

模型	公开数据集 RMSE	公开数据集 MAE	鱼跃数据集 RMSE	鱼跃数据集 MAE
TCN-GRU	0.594	0.483	0.582	0.470
TCN-LSTM	0.607	0.491	0.592	0.483
LSTM	0.793	0.563	0.788	0.526
ARIMA	0.662	0.512	0.670	0.522
RNN	0.896	0.591	0.906	0.647
T-GA	0.571	0.430	0.546	0.419

在公开数据集上,T-GA 模型的 RMSE 为 0.571,相较于 TCN-GRU 和 TCN-LSTM 模型实现了 0.023 和 0.036 的优化,同时分别比 LSTM 和 ARIMA 提升了 0.222 和 0.091。RNN 在血糖预测任务中性能最差。在 MAE 指标方面,T-GA 同样取得了更好的效果,分别比 TCN-GRU 和 TCN-LSTM 优化了 0.053 和 0.061,分别优于 LSTM 和 ARIMA 的 MAE 值 0.133 和 0.082。在鱼跃数据集上,各模型的预测效果与上述相似,T-GA 模型取得了更好的性能表现,RMSE 和 MAE 分别是 0.546 和 0.419,均优于其他模型的评估结果。

相比 RNN 在处理长序列时可能会发生梯度消失^[41]从而难以有效捕捉长距离时序依赖关系的不足,GRU 突破了反向传播中梯度问题的局限,而 LSTM 相较于 GRU 训练过程更慢。ARIMA 模型更适用于处理平稳时间序列数据中的线性关系。同时,在 TCN-GRU 模型中,TCN 利用扩张因果卷积模拟时间序列数据中的复杂依赖关系,且残差结构加速了模型收敛速度,在 GRU 中应用注意力机制后,模型能更好地权衡不同时间点的重要性,提高对血糖模式细微变化的敏感性。可以得出结论,T-GA 模型在时序建模的血糖预测任务中展现了显著的优势,主要得益于其综合了麻雀搜索算法、TCN、GRU 以及注意力机制等技术的优点。

为进一步验证 SSA 在血糖预测任务中的有效性,使用 SSA 同时优化 T-GA 和上述 LSTM 等对比模型,并评估优化后的模型在血糖预测任务中的性能。表 6 展示了通过 SSA 优化后的不同模型的 RMSE 和 MAE 结果。

表 6 SSA 优化不同深度学习模型的血糖预测实验结果对比

Table 6 Comparison of blood glucose prediction experiment results of different deep learning models after SSA optimization

模型	公开数据集 RMSE	公开数据集 MAE	鱼跃数据集 RMSE	鱼跃数据集 MAE
S-TCN-GRU	0.579	0.445	0.561	0.422
S-TCN-LSTM	0.580	0.465	0.579	0.469
S-LSTM	0.763	0.514	0.744	0.501
S-ARIMA	0.585	0.483	0.626	0.497
S-RNN	0.766	0.536	0.872	0.608
S-T-GA	0.552	0.402	0.531	0.388

经过采用 SSA 进行超参数优化后,公开数据集和鱼跃数据集上的血糖预测模型性能均得到显著提升。在公开数据集上,S-T-GA 模型的 RMSE 为 0.552,MAE 为 0.402,相较于未优化前分别提升了 0.019 和 0.028;而在鱼跃数据集上,S-T-GA 模型的 RMSE 和 MAE 分别为 0.531 和 0.388,进一步提升了 0.015 和 0.031。同时,经过 SSA 优化后的 LSTM 和 ARIMA 模型在两个数据集上的表现也有显著改善,S-TCN-GRU 和 S-TCN-LSTM 模型在两个数据集上预测效果次之。S-RNN 在两个数据集上的效果相对较差。

SSA 是一种高效、全局搜索能力强且易调整的优化算法,能够灵活地适应多种深度学习模型的优化需求,上述结果验证了 SSA 优化算法在改进血糖预测任务中的有效性,提高了模型的性能。

公开数据集以第 5 名参与者为例,鱼跃数据集以其中一名患者的 11 月 1 日数据为例,分别展示其前 50 条血糖预测结果,并将不同的神经网络模型在血糖预测结果上的差异进行了对比。图 6 和图 7 分别展示了两种数据集血糖预测的具体情况,其中横坐标表示时间,纵坐标表示血糖值。

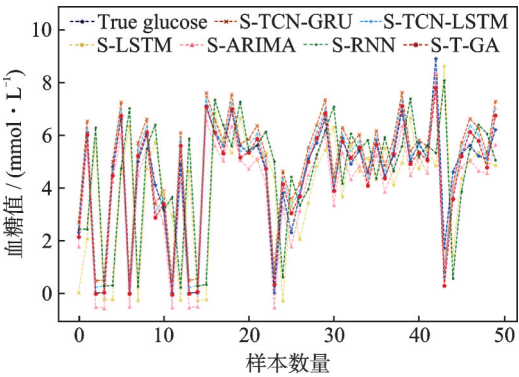


图 6 不同神经网络在公开数据集上的实验结果对比
Fig.6 Comparison of experimental results of different neural networks on public dataset

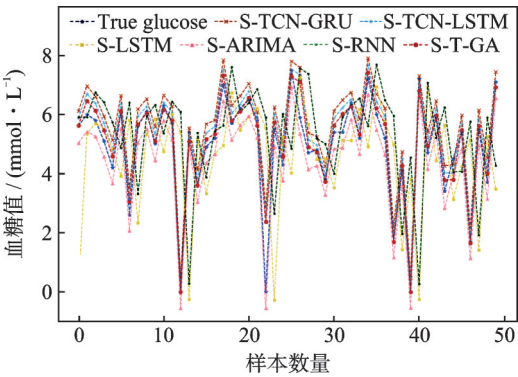


图 7 不同神经网络在鱼跃数据集上的实验结果对比
Fig.7 Comparison of experimental results of different neural networks on YuWell dataset

图 6、7 中的蓝色虚线代表实际血糖值,红色虚线代表 S-T-GA 混合模型的预测结果。观察图 6 所知,在公开数据集的血糖拟合结果中,S-T-GA 模型在大多数情况下与真实血糖值非常接近,呈现出更好的拟合和预测能力,S-TCN-GRU、S-TCN-LSTM 和 S-ARIMA 效果次之,S-LSTM 的初始拟合效果较差,S-RNN 在血糖预测任务中表现最为不理想。如图 7 所示,S-T-GA 和其他对比模型在鱼跃数据集上的拟合效果与公开数据集的结果相似。

上述结果表明 LSTM 模型^[37]在捕获长时间序列中的血糖模式方面表现不佳,而 GRU 相对于传统的 RNN^[41]具有更高的灵活性,能够选择性地更新信息,更有效地捕获和保留输入序列中的相关模式。此外,ARIMA 不擅长非线性拟合,且缺乏内置的门控机制^[42],而 GRU 中的门控机制更有助于学习和表征时间依赖性。同时,通过 TCN 的扩展因果卷积和残差结构、GRU 的门控机制以及注意力机制的有机结合。本文提出的 S-T-GA 模型在血糖预测任务中脱颖而出,成功地捕捉到了血糖波动的复杂模式,取得了良好的预测效果。

4 结束语

在需求响应迅速的医疗应用中,现有血糖预测模型的训练过程存在较大的时间开销,且多数研究倾向于采用单一神经网络模型。为解决这些问题,本文提出了一种基于麻雀搜索算法的混合神经网络模型,并将其应用于实际血糖预测任务。首先,通过充分利用 TCN 模型的并行计算特性,加速了模型训练和推断过程。TCN 通过扩张因果卷积捕获血糖数据在不同尺度下的时序特征,而其残差结构有效避免了梯度消失或梯度爆炸问题。然后,通过基于注意力机制的 GRU 模型对血糖数据特征进行非线性拟合,注意力机制在 GRU 快速建模的同时增强了对不同时间点关键信息的捕获,更好地聚焦长期时序依赖关系。最后,通过 SSA 对 TCN 和 GRU 模型进行超参数优化,将得到优化的超参数组合用于血糖预测,最终实现了该混合神经网络模型在血糖预测任务中的应用。同时,通过与其他深度学习模型比较,本文提出的 S-T-GA 模型展现出更好的预测精度。然而,本研究中的模型组合是基于人工选择而非智能机制筛选,因此希望在未来的研究工作中,可进一步探讨多种神经网络模型的灵活组合结构,通过建立一个基学习器库,并结合麻雀搜索等群体智能优化算法实现基学习器的自动组合,从而在血糖预测任务中进一步提升预测精度。

参考文献:

- [1] 王尚才. 饮食结构与糖尿病发生发展的相关性研究[J]. 中国食物与营养, 2021, 27(2): 68-71.
WANG Shangcai. The correlation between diet structure and the occurrence and development of diabetes mellitus[J]. Food and Nutrition in China, 2021, 27(2): 68-71.
- [2] LIN L, QIN S, WU Z, et al. Facial expression recognition in video sequence based on LBP feature and GRU[C]//Proceedings of the 2021 5th International Conference on Video and Image Processing. Hayward, CA, USA: ACM, 2022: 38-43.
- [3] WANG F, YAN J, XU H. A special multivariate polynomial model for diabetes prediction and analysis[C]//Proceedings of the 3rd International Symposium on Artificial Intelligence for Medicine Sciences. Amsterdam, Netherlands: ACM, 2022: 302-306.
- [4] 周思思, 李勇, 郭钊秀, 等. 考虑时序特征提取与双重注意力融合的 TCN 超短期负荷预测[J]. 电力系统自动化, 2023, 47(18): 193-205.
ZHOU Sisi, LI Yong, GUO Yixiu, et al. Ultra-short-term load forecasting based on temporal convolutional network considering temporal feature extraction and dual attention fusion[J]. Automation of Electric Power Systems, 2023, 47(18): 193-205.
- [5] ALOTAIBI R. ARIMA model for stock market prediction[C]//Proceedings of the 2022 8th International Conference on Computer Technology Applications. Vienna, Austria: ACM, 2022: 1-4.
- [6] CHENG Y, ZHANG F, HU G, et al. Block popularity prediction for multimedia storage systems using spatial-temporal-sequential neural networks[C]//Proceedings of the 29th ACM International Conference on Multimedia. [S.l.]: Association for Computing Machinery, 2021: 3390-3398.
- [7] TAMAKI R, FUJIMOTO M, SUWA H, et al. Data analysis for developing blood glucose level control system[C]//Proceedings of International Conference on Distributed Computing and Networking. Nara, Japan: ACM, 2021: 104-109.
- [8] 闫洪举. 基于深度学习的金融时间序列数据集成预测[J]. 统计与信息论坛, 2020, 35(4): 33-41.
YAN Hongju. Integrated prediction of financial time series based on deep learning[J]. Journal of Statistics and Information, 2020, 35(4): 33-41.

- [9] CHAUDHURI T, WU M, ZHANG Y, et al. An attention-based deep sequential GRU model for sensor drift compensation[J]. IEEE Sensors Journal, 2020, 21(6): 7908-7917.
- [10] 蒋泉哲, 任冠青, 周烽, 等. 基于多任务学习时间卷积网络的路网短时交通速度预测[J]. 中国高科技, 2022(24): 56-58.
JIANG Xiaozhe, REN Guanqing, ZHOU Feng, et al. Short-term traffic speed prediction of road network with multi-task learning temporal convolutional network[J]. China High and New Technology, 2022(24): 56-58.
- [11] SEHOVAC L, GROLINGER K. Deep learning for load forecasting: Sequence to sequence recurrent neural networks with attention[J]. IEEE Access, 2020, 8: 36411-36426.
- [12] RABBY M F, TU Y, HOSSEN M I, et al. Stacked LSTM based deep recurrent neural network with Kalman smoothing for blood glucose prediction[J]. BMC Medical Informatics and Decision Making, 2021, 21: 1-15.
- [13] YANG J, LI L, SHI Y, et al. An ARIMA model with adaptive orders for predicting blood glucose concentrations and hypoglycemia[J]. IEEE Journal of Biomedical and Health Informatics, 2018, 23(3): 1251-1260.
- [14] MARTINSSON J, SCHLIEP A, ELIASSON B, et al. Blood glucose prediction with variance estimation using recurrent neural networks[J]. Journal of Healthcare Informatics Research, 2020, 4: 1-18.
- [15] PENG X, LI Q, WANG Y, et al. High-precision blood glucose prediction and hypoglycemia warning based on the LSTM-GRU model[J]. Trends in Computer Science and Information Technology, 2022, 7(3): 74-80.
- [16] LIU J W, LIU J W, LUO X L. Research progress in attention mechanism in deep learning[J]. Chinese Journal of Engineering, 2021, 43(11): 1499-1511.
- [17] 高凯悦, 牟莉, 张英博. PCC-BiLSTM-GRU-Attention组合模型预测方法[J]. 计算机系统应用, 2022, 31(7): 365-371.
GAO Kaiyue, MOU Li, ZHANG Yingbo. Prediction method of PCC-BiLSTM-GRU-Attention combined model[J]. Computer Systems & Applications, 2022, 31(7): 365-371.
- [18] YAN C, TU Y, WANG X, et al. STAT: Spatial-temporal attention mechanism for video captioning[J]. IEEE Transactions on Multimedia, 2019, 22(1): 229-241.
- [19] RAN X, SHAN Z, FANG Y, et al. An LSTM-based method with attention mechanism for travel time prediction[J]. Sensors, 2019, 19(4): 861.
- [20] ZHANG X, LIU S, ZHENG X. Stock price movement prediction based on a deep factorization machine and the attention mechanism[J]. Mathematics, 2021, 9(8): 800.
- [21] CHEN Y, PENG G, ZHU Z, et al. A novel deep learning method based on attention mechanism for bearing remaining useful life prediction[J]. Applied Soft Computing, 2020, 86: 105919.
- [22] 滕建丽, 容芷君, 许莹, 等. 基于GRU网络的血糖预测方法研究[J]. 计算机应用与软件, 2020, 37(10): 107-112.
TENG Jianli, RONG Zhijun, XU Ying, et al. Blood glucose prediction method based on GRU[J]. Computer Applications and Software, 2020, 37(10): 107-112.
- [23] 李雅丽, 王淑琴, 陈倩茹, 等. 若干新型群智能优化算法的对比研究[J]. 计算机工程与应用, 2020, 56(22): 1-12.
LI Yali, WANG Shuqin, CHEN Qianru, et al. Comparative study of several new swarm intelligence optimization algorithms[J]. Computer Engineering and Applications, 2020, 56(22): 1-12.
- [24] 高兵, 郑雅, 秦静, 等. 基于麻雀搜索算法和改进粒子群优化算法的网络入侵检测算法[J]. 计算机应用, 2022, 42(4): 1201-1206.
GAO Bing, ZHENG Ya, QIN Jing, et al. Network intrusion detection algorithm based on sparrow search algorithm and improved particle swarm optimization algorithm[J]. Journal of Computer Applications, 2022, 42(4): 1201-1206.
- [25] WANG W, TONG M, YU M. Blood glucose prediction with VMD and LSTM optimized by improved particle swarm optimization[J]. IEEE Access, 2020, 8: 217908-217916.
- [26] AN G, JIANG Z, CHEN L, et al. Ultra short-term wind power forecasting based on sparrow search algorithm optimization deep extreme learning machine[J]. Sustainability, 2021, 13(18): 10453.
- [27] BUI Q T, NGUYEN Q H, NGUYEN X L, et al. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping[J]. Journal of Hydrology, 2020, 581: 124379.
- [28] CHEN M, SHI H, WU J. Research on transformer fault diagnosis based on sparrow algorithm optimization probabilistic neural network[C]//Proceedings of the 6th International Conference on Intelligent Information Processing. Bucharest, Romania: ACM, 2021: 254-259.
- [29] 尹德鑫, 张达敏, 张琳娜, 等. 认知工业物联网中基于麻雀搜索算法的频谱分配策略[J]. 数据采集与处理, 2022, 37(2):

371-382.

YIN Dexin, ZHANG Damin, ZHANG Linna, et al. Spectrum allocation strategy based on sparrow algorithm in cognitive industrial internet of things[J]. *Journal of Data Acquisition and Processing*, 2022, 37(2): 371-382.

- [30] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法[J]. *北京航空航天大学学报*, 2021, 47(8): 1712-1720.
LYU Xin, MU Xiaodong, ZHANG Jun, et al. Chaos sparrow search optimization algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(8): 1712-1720.
- [31] ZHANG D, XU Y. Sparrow search algorithm-optimized long short-term memory model for network latency prediction[C]// *Proceedings of the 2022 5th International Conference on Algorithms, Computing and Artificial Intelligence*. Sanya, China: ACM, 2023.
- [32] LI C, GUO Q, SHAO L, et al. Research on short-term load forecasting based on optimized GRU neural network[J]. *Electronics*, 2022, 11(22): 3834.
- [33] 曲扬, 陈波, 李佳鹏, 等. 基于滑动平均法的施工过程监测数据解耦方法[J]. *建筑施工*, 2021, 43(6): 1163-1165.
QU Yang, CHEN Bo, LI Jiapeng, et al. Decoupling method of construction process monitoring data based on moving average method[J]. *Building Construction*, 2021, 43(6): 1163-1165.
- [34] ZHU L, LI M, METAWA N. Financial risk evaluation Z-score model for intelligent IoT-based enterprises[J]. *Information Processing & Management*, 2021, 58(6): 102692.
- [35] HENDERI H, WAHYUNINGSIH T, RAHWANTO E. Comparison of min-max normalization and Z-score normalization in the K-nearest neighbor (kNN) algorithm to test the accuracy of types of breast cancer[J]. *International Journal of Informatics and Information Systems*, 2021, 4(1): 13-20.
- [36] LINDSAY G W. Convolutional neural networks as a model of the visual system: Past, present, and future[J]. *Journal of Cognitive Neuroscience*, 2021, 33(10): 2017-2031.
- [37] 杨海涛, 王华朋, 牛瑾琳, 等. 融合 LSTM-GRU 网络的语音逻辑访问攻击检测[J]. *数据采集与处理*, 2022, 37(2): 396-404.
YANG Haitao, WANG Huapeng, NIU Jinlin, et al. Logical access attack audio detection based on LSTM-GRU[J]. *Journal of Data Acquisition and Processing*, 2022, 37(2): 396-404.
- [38] JIANG Z. An attention GRU-XGBoost model for stock market prediction strategies[C]// *Proceedings of the 4th International Conference on Advanced Information Science and System*. Sanya, China: ACM, 2023.
- [39] 马佳成, 王晓霞, 杨迪. 基于 Attention 机制的 TCN-LSTM 非侵入式负荷分解[J]. *电力信息与通信技术*, 2023, 21(8): 43-51.
MA Jiacheng, WANG Xiaoxia, YANG Di. Non-intrusive load decomposition based on TCN-LSTM model with Attention mechanism[J]. *Electric Power Information and Communication Technology*, 2023, 21(8): 43-51.
- [40] 刘可真, 阮俊泉, 赵现平, 等. 基于麻雀搜索优化的 Attention-GRU 短期负荷预测方法[J]. *电力系统及其自动化学报*, 2022, 34(4): 99-106.
LIU Kezhen, RUAN Junxiao, ZHAO Xianping, et al. Short-term load forecasting method based on sparrow search optimized Attention-GRU[J]. *Proceedings of the CSU-EPSA*, 2022, 34(4): 99-106.
- [41] VISHNUMOHANM S, MENON V. Measuring viscosity of fluids: A deep learning approach using a CNN-RNN architecture [C]// *Proceedings of the 1st International Conference on AI-ML Systems*. Bangalore, India: ACM, 2021: 1-5.
- [42] 石庆研, 岳聚财, 韩萍, 等. 基于 LSTM-ARIMA 模型的短期航班飞行轨迹预测[J]. *信号处理*, 2019, 35(12): 2000-2009.
SHI Qingyan, YUE Jucai, HAN Ping, et al. Short-term flight trajectory prediction based on LSTM-ARIMA model[J]. *Journal of Signal Processing*, 2019, 35(12): 2000-2009.

作者简介:



徐鹤(1985-), 男, 博士, 教授, CCF 高级会员, 研究方向: 人工智能与大数据、物联网技术, E-mail: xuhe@njupt.edu.cn.



许硕洋(1999-), 女, 硕士研究生, 研究方向: 人工智能与大数据、深度学习, E-mail: 1222045830@njupt.edu.cn.



季一木(1978-), 通信作者, 男, 博士, 教授, 博士生导师, CCF 会员, 研究方向: 高性能计算、大数据和人工智能, E-mail: jiym@njupt.edu.cn.

(编辑: 王静)