

基于多头注意力机制的多模态帕金森病 安全检测系统

季培琛^{1,2}, 李晨¹

(1. 徐州市中医院, 江苏 徐州 221000; 2. 徐州医科大学, 江苏 徐州 221000)

摘要: 在实际的帕金森病远程诊断过程中, 应用单模态数据检测帕金森病存在误诊率较高的问题, 且远程诊断的安全性问题突出; 为提高帕金森病远程诊断准确率与安全性, 设计一种具有隐私保护功能的帕金森病多模态安全远程辅助检测系统; 使用帕金森病语音和步态双模态数据, 在传统卷积神经网络后融合多头注意力机制与多层感知机, 有效提高模型的特征提取、融合与识别能力; 为了保证数据传输过程的安全性, 使用基于余弦混沌的差分隐私加噪方式扰动随机拆分的数据编号, 提高帕金森病数据传输安全性; 通过两模态消融实验和对比实验结果表明, 提出的基于多头注意力机制的帕金森病多模态远程检测模型实际测试准确率达到 0.913, 且模型的各项评估指标和收敛速度等均高于传统模型, 具备良好的帕金森病智能辅助检测效果, 能够满足帕金森病早期智能安全筛查与诊断需求。

关键词: 帕金森病; 多头注意力机制; 余弦混沌; 差分隐私; 多模态远程检测

Multimodal Safety Detection System for Parkinson's Disease Based on Multi-head Attention Mechanism

Ji Peichen^{1,2}, Li Chen¹

(1. Xuzhou City Hospital of TCM, Xuzhou 221000, China;

2. Xuzhou Medical University, Xuzhou 221000, China)

Abstract: During the actual remote diagnosis process of Parkinson's disease, it has a high misdiagnosis rate that the single-mode data is applied to detect Parkinson's disease, and the remote diagnosis has prominent security. To improve the accuracy and security of remote diagnosis of Parkinson's disease, a multimodal secure remote auxiliary detection system with a privacy protection function was designed. The dual-modal data of Parkinson's disease speech and gait were used to integrate the multi-head attention mechanism with multi-layer perceptron after traditional convolutional neural networks, which effectively improved the feature extraction, fusion, and recognition ability of the model. To ensure the security of the data transmission process, the differential privacy and noise method based on cosine chaos is used to disturb the randomly separated data numbers to improve the data transmission security of Parkinson's disease. The two-mode ablation experiment and comparison experiment results show that the actual test accuracy of the proposed multi-mode remote detection model of Parkinson's disease based on the multi-head attention mechanism reaches by 0.913. The evaluation indicators and convergence speed of the model are higher than that of the traditional model, which has a good effect on intelligently assisted detection of Parkinson's disease. It can meet the needs of early innovative safety screening and diagnosis of Parkinson's disease.

Keywords: parkinson's disease; multi-head attention mechanism; cosine chaos; differential privacy; multi-modal remote detection

0 引言

帕金森病 (PD, parkinson's disease) 是一种常见的神经系统退行性疾病, 目前该病发病机制尚未明确, 主要受到遗传、环境、年龄老化以及氧化应激等诸多因素影响^[1-2]。据统计, 年龄在 45 岁以上和 65 岁以上人群中, PD 的发病率分别为 0.4% 和 1.7%^[3], 预计到 2030 年, 我国

PD 患者将达 500 万, 在全球排名第一, 约占世界 50%。通过众多病历的长期跟踪和数据统计结果显示, 随着 PD 患者病情发展以及年龄的增长, 患者各项身体机能将逐渐退化、行动也受到严重限制, 导致其 6 年死亡率高达 66%, 要明显高于慢性心力衰竭 (50.9%)、慢性阻塞性肺病 (44.7%)、残血性心脏病 (32.5%)、中风或短暂性脑缺血发作 (52.5%) 等疾病。而 PD 缺少明确的病理机制, 早期

收稿日期: 2024-01-22; 修回日期: 2024-02-04。

基金项目: 国家自然科学基金(62003291)。

作者简介: 季培琛(1984-), 男, 硕士。

引用格式: 季培琛, 李晨. 基于多头注意力机制的多模态帕金森病安全检测系统[J]. 计算机测量与控制, 2024, 32(3): 138-145.

症状隐匿, 并存在非运动症状和相似神经系统疾病症状的干扰, 导致 PD 早期诊断极为困难。

在大量的临床试验中^[4-6], 语音功能障碍和步态特征是 PD 患者临床表现中非常典型的症状。PD 患者在语音特征上多表现为语速慢、停顿增多、音质颤抖及刺耳等症状, 在步态特征上多表现为快速小碎步、拖把步、平衡性差等症状。众多学者利用 PD 患者与正常人的语音和步态特征差异, 使用智能算法, 开展了大量基于语音和步态数据的 PD 辅助诊断研究。例如, Little 等^[7]利用模式识别方法, 对基于语音障碍的 PD 诊断进行分析, 并建立了首个 PD 语音障碍数据集; Wroge^[8]等使用深度神经网络监督分类算法, 结合语音数据, 完成 PD 的智能诊断, 峰值准确率为 85%; 朱家英^[9]等提出了基于多尺度特征和动态注意力机制的多模态循环融合模型, 实现了对 PD 患者的识别与检测。

而在实际研究中发现, 用于 PD 辅助检测的语音数据中, 包含共振峰频率、音调、重音等可唯一识别个体的声纹特征^[10-11]。同时, 步态数据中也包含了步频、步长、步态周期、膝盖弯曲角度等可以唯一识别特定个体的运动学特征和姿势特征^[12-13]。而在已有的相关研究中, 众多学者往往忽略了对 PD 患者隐私安全的保护, 极易在数据传输过程中发生隐私泄露, 且很难实现 PD 的多模态辅助诊断准确性与隐私安全的动态平衡。为此, 本文设计了一种鲁棒性高、成本低且操作便捷的基于多头注意力机制^[14-15]的帕金森病多模态安全远程诊疗模型, 通过语音和步态两模态数据特征的提取和识别, 使 PD 诊断结果更加精准, 也更具临

床参考价值。同时引入基于余弦混沌的差分隐私噪声扰动方法, 实现了对 PD 数据传输过程的安全保护, 为 PD 早期远程辅助诊断和 PD 诊断临床决策支持提供了支撑。本文主要创新和贡献如下:

- 1) 针对传统 PD 检测模型训练和测试数据存在的特征固定、模态单一问题, 提出了基于多头注意力机制的两模态特征融合与识别模型, 避免了单一模态数据噪声干扰、数据规模小导致的检测识别准确度低等问题, 实现了基于两模态数据特征的 PD 智能检测。
- 2) 针对现有 PD 智能检测识别研究多忽略数据主体隐私安全的问题, 设计了一种基于余弦混沌的差分隐私噪声扰动方式, 通过扰动随机拆分的数据编号, 保证数据传输至系统智能检测识别模块的传输过程安全性, 实现了 PD 检测准确率和隐私安全的动态平衡。
- 3) 设计了基于多头注意力机制的多模态特征融合方法, 在特征融合阶段, 通过挖掘 PD 语音特征与步态数据特征的内在相关性, 提高了模型的疾病表征能力, 并具有良好的多模态特征融合扩展性, 可满足更高模态特征融合与识别需求。

1 系统设计

基于多头注意力机制的多模态帕金森病安全检测系统整体框架如图 1 所示。

1.1 系统结构及原理

系统主要由 3 部分组成, 第一层为数据采集层, 主要

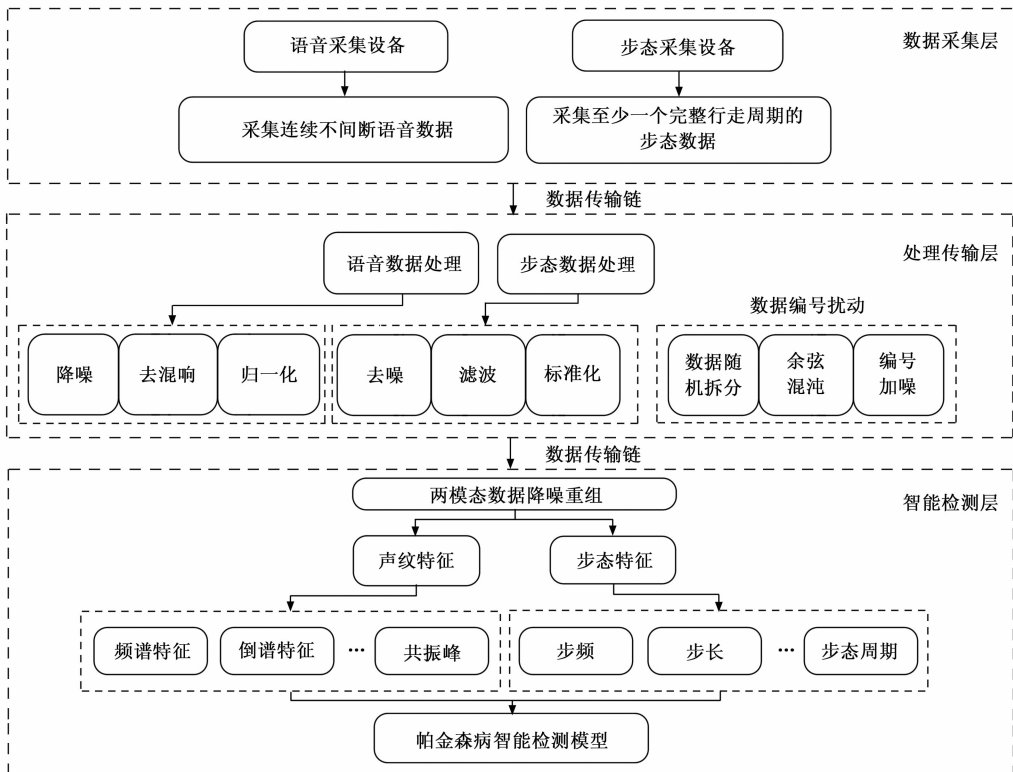


图 1 基于多头注意力机制的多模态帕金森病安全检测系统框架

借助语音录入设备（如智能手机、树莓派、录音设备等）和步态数据采集设备（如摄像机、运动相机、智能平板等）来完成受试者语音数据和步态数据的采集，其中语音数据应确保为一段连续不间断的语音数据，便于更好记录受试者音色、音调等声纹特征的微小变化；步态数据应为连续不间断的视频数据，记录受试者完整的行走周期信息，以便于分析受试者步长、步频、步态等步态特征。第二层为数据处理和传输层，该层主要将采集后的数据进行清洗和处理，并使用基于余弦混沌的数据随机拆分编号加噪方式，打乱数据顺序，保证两模态数据上传过程安全。第三层为 PD 智能诊断层，接收到上传的数据后，进行数据编号逆向降噪，得到完整的两模态语音和步态数据特征，并将降噪和特征提取后的语音声纹特征和步态特征作为输入数据，使用融合多头注意力机制的卷积神经网络完成帕金森病的安全检测。

1.2 系统设计目标分析

为确保基于多头注意力机制的帕金森病多模态安全远程检测系统能够高效、精准且安全的实现 PD 远程辅助检测，为医生诊断 PD 提供临床决策支持，该系统设计应实现以下几个目标。

1) 鲁棒性：指的是系统应能有效处理各种异常数据，如设备故障、信号干扰等，避免因个别异常数据导致系统崩溃或数据丢失；同时应保证网络、容错及恢复的鲁棒性，确保系统在网络故障、通信延迟等情况下，仍能保持多模态诊疗数据传输的稳定性和可靠性。在本研究中，要求系统提供更全面和准确地诊断信息，能够通过多模态数据融合降低单一数据源的误差，同时具有故障检测和恢复机制，当检测到异常或错误时，系统能够自动调整或切换到备用方案，确保服务的连续性和稳定性，从而提高系统鲁棒性。

2) 安全性：指的是系统应采用严格的数据加密和访问控制机制，保证患者数据传输和存储过程的机密性和完整性；同时融合多因素身份验证和细粒度授权机制，保证只有经过授权的人员能够访问敏感数据或执行相关操作。在本系统中用于远程诊断 PD 的语音、步态等多模态数据中包含患者大量的隐私信息，在数据采集完成后，经过安全处理后上传智能辅助诊断模块；同时系统的应用需要建立明确的数据安全、网络保护与访问控制方案，保证 PD 多模态远程诊疗系统的硬件设备的物理安全，确保系统使用全流程可追溯和审计，避免出现隐私泄露，确保系统以及患者隐私安全性。

3) 准确性：指的是系统在识别和判断疾病时的准确程度，是评估系统性能的重要评估指标之一。在该系统中，利用多头注意力机制，能够自动提取识别测试人员上传的与帕金森病相关的多模态特征，减少人为因素和主观判断对诊断结果的影响。同时借助深度学习技术，在借助大量训练数据进行模型训练的基础上，迭代和优化 PD 远程辅助检测算法模型，并通过临床验证完成系统的多轮更新和

优化。

4) 可扩展性：指的是系统在面对新型疾病和数据时，能够适应并快速进行自适应的扩展和改进，以适应新的需求。在本研究中，系统设计采用模块化方法，确保数据收集、处理、分析以及可视化板块的相互独立，保证各模块能够独立运行；同时系统应具备计算资源的动态分配和弹性配置能力，并满足持续开发、功能更新与迭代的扩展能力，满足多模态与跨模态检测 PD 的需求。

5) 易用性：指的是产品、系统或服务对用户而言的易于理解、学习和操作程度。它包括界面设计的友好程度、操作的直观性以及用户完成任务的效率和满意度。在本文中，系统的设计应简单易用，方便测试人员、技术人员和医务人员的使用。同时界面设计简单大方，操作流程简单易懂，并提供丰富的交互反馈和引导功能，能够为医生临床诊断 PD 提供决策支持。

2 系统软件设计

该系统设计主要通过识别早期 PD 患者在语音和步态特征上异于常人的表现或障碍，来实现早期 PD 的安全智能检测。为提高 PD 智能检测精确度，降低单一模态数据辅助检测存在的噪声干扰、数据稀疏问题影响，使用语音、步态两模态数据辅助检测早期 PD。为保证数据传输过程安全性，将采集的数据进行随机分组和编号，并使用基于余弦混沌的差分噪声添加方式，扰动数据编号，防止数据攻击和重组导致的数据隐私披露。数据上传后，分别进行语音和步态数据特征提取，并使用多头注意力机制完成两模态数据特征融合，特征融合后作为输入数据输入 PD 智能检测模型中，最终完成 PD 的智能检测。系统关键技术及算法模型设计如下。

2.1 相关技术简介

2.1.1 多头注意力机制

多头注意力机制 (MHA, multi-head attention) 是神经网络中的一种注意力机制。MHA 能够使诊断算法模型在处理输入数据时从多个视角上关注不同模态数据的特征子集，帮助更加全面地理解和聚焦于疾病诊断的关键特征信息。通过 MHA 的应用，可以有效提升模型整合生理信号、文本、图像、视频等数据特征的能力，进一步提高复杂疾病的诊断准确性。目前 MHA 已广泛应用于肿瘤、神经系统疾病、心血管疾病等的辅助诊断和研究中，并取得了较好的研究效果。

在多头注意力机制中，主要分为 3 个步骤，计算注意力权重，进行多个头的计算，以及进行输出层的计算并拼接。在 MHA 中主要包含查询向量 $Q \in R^d$ 、键向量 $K \in R^d$ 和值向量 $V \in R^d$ 。其中，可学习的参数包括 W^Q, W^K, W^V ，以及代表注意力汇聚的函数 $Multihead_{fusion}$ 。其中 $Multihead_{fusion}$ 可以是加性注意力和缩放点积注意力。多头注意力的输出需要经过另一个线性变换，它对应着 h 个头连接后的结果，因此其可学习参数是 W^O 。

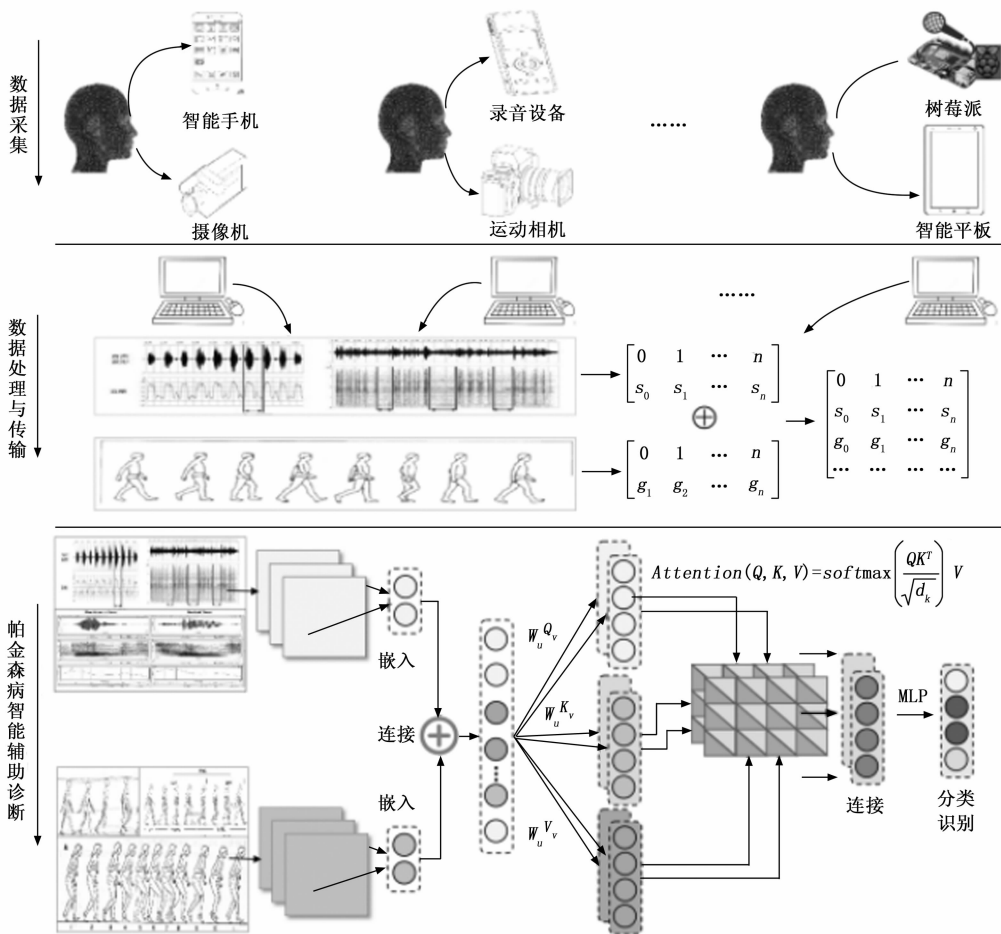


图 2 基于多头注意力机制的多模态 PD 安全检测系统技术框架

2.1.2 差分隐私

差分隐私保护通过对数据加噪掩盖原始数据的真实值, 确保攻击人员无法结合背景知识等推断出相关数据, 从而达到隐私保护的目的。研究人员可以根据研究场景和需求自适应的设计噪声添加方式, 以达到最大化准确率和安全性的目标。差分隐私的数学定义如下:

对于任意相邻的数据集 $D, D' \in Z$, 给定一个随机算法 $f: Z \rightarrow R$, 和任意输出结果 $S \subseteq R$, 则定义以下不等式:

$$\max_s \left[\ln \frac{Pr[f(D) \in S] - \xi}{Pr[f(D') \in S]} \right] \leq \epsilon \quad (1)$$

若不等式 (1) 成立, 则成算法 f 满足差分隐私定义^[16]。其中 ϵ 为隐私预算, 表示可以提供的随机化算法的保护级别, 当 ϵ 越小时, 表示隐私保护强度越强, 即要求添加的噪声越大; 反之 ϵ 越大, 表示隐私保护强度越小, 即要求添加的噪声越小。 ξ 为一个非零实数, 通常是一个很小的数值, 表示不满足上述不等式的概率。

2.2 基于多头注意力机制的多模态 PD 检测模型设计

由于早期 PD 患者在音调、音量、语速以及音质等语音特征异于正常人的表现^[17-18], 目前更多的是借助单一模态的语音信号开展 PD 智能辅助诊断研究。但由于语音信号易

受到语音采集设备、外部环境噪音等干扰, 导致基于单模态语音信号识别 PD 的辅助诊断结果存在不稳定和误差较大等局限。因此, 为提升 PD 智能辅助诊断准确率, 借助多模态数据开展 PD 智能辅助诊断是可行之路。

在现有研究中, 使用卷积神经网络 (CNN, convolutional neural networks) 在全连接层进行双峰数据融合, 来检测和识别早期 PD 是一种常见的方法^[19]。但这种融合方法不利于多模态数据特征间的相关性信息挖掘和使用。为解决此问题, 提出了一种融合多头注意力机制的 MHA-CNN, 来获取语音、步态等多模态数据间的相关信息权重, 以更好的提取和融合高维特征表示。

在多头注意力机制中, 引入多个注意力头, 将输入的语音和步态数据分成多份, 每个注意力头独立地学习并关注不同的语义信息, 有效增强模型的表达能力与性能。基于多头注意力机制的多模态 PD 智能检测识别模型设计如下。

假设输入 2-dimension 模态为模型的输入数据。在 MHA-CNN 完成特征提取后, 使用 $\{X^1, X^2\}$ 分别表示语音和步态数据, 使用 $\{d_1, d_2\}$ 分别表示语音和步态数据嵌入, 可得到:

$$X^S = \{X_1^S, X_2^S, \dots, X_N^S\} \in R^{N \times d_s} \quad (2)$$

$$X^G = \{X_1^G, X_2^G, \dots, X_N^G\} \in R^{N \times d_g} \quad (3)$$

在 concat 层将语音和步态数据特征向量进行拼接:

$$X^{\text{concat}} = \{X_1^S + X_1^G, X_2^S + X_2^G, \dots, X_N^S + X_N^G\} \in R^{N \times (d_s + d_g)} \quad (4)$$

在全连接层, 将嵌入的语音和步态数据进行融合, 定义 d_{concat} 和 F_{concat} :

$$d_{\text{concat}} = d_s + d_g \quad (5)$$

$$F_{\text{concat}} = W_{\text{concat}} F_{\text{concat}} + b \quad (6)$$

其中: $F_{\text{concat}} \in R^{N \times d_{\text{concat}}}$.

在多头注意力机制中, 每个注意力头独立地学习并关注不同的语义信息, 通过计算查询向量和键向量的相似度来获得注意力权重值, 进而根据权重对值向量进行加权求和, 得到最终的输出表示。

定义语音信号与步态数据间的相似度关系为 r , 不同 PD 患者语音和步态数据间的 r 通过计算公式可表示为:

$$r = \frac{QK^T}{\sqrt{d_K}} \quad (7)$$

使用 softmax 函数计算语音和步态数据两模态特征权重值为:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (8)$$

其中: Q 、 K 、 V 的可表示为:

$$Q = [Q_1, Q_2, \dots, Q_N] \in R^{N \times d_{\text{query}}} \quad (9)$$

$$K = [K_1, K_2, \dots, K_N] \in R^{N \times d_{\text{key}}} \quad (10)$$

$$V = [V_1, V_2, \dots, V_N] \in R^{N \times d_{\text{value}}} \quad (11)$$

研究中使用多头注意力机制改进 CNN 网络, 增强 MHA-CNN 模型关注语音和步态两模态的能力, 使分区不同的头相互集中, 同时通过将输入特性划分为单独的分区来为其添加子空间, 便于从语音和步态两模态数据特征子空间学习到更多不同信息。其中基于两模态数据的头部注意力计算公式为:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (12)$$

经过独立计算头部注意力, 将结果输出后连接, 用来获取所有子空间的特征信息, 并反馈到线性投影中获得最终的两模态特征融合模型维度, 计算公式如下:

$$\text{Multihead}_{\text{fusion}} = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h)W^O \quad (13)$$

两模态数据特征提取融合后, 进一步使用多层感知机 MLP^[20] 按照标记的 PD 患者语音和步态数据特征进行分类, 并返回预测结果。整体过程如下:

算法 1: 基于 MHA 的两模态数据融合模型

输入: 数据集 $D = \{(X^S \in R^{N \times d_s}, X^G \in R^{N \times d_g}), Y\}$, 注意力头数为 H , 学习率为 η

输出: PD 检测识别结果 Dr

初始化模型参数 W, b

for each round $t=1, 2, \dots, n$ do

Step1: 两模态特征提取

从 X^S 和 X^G 中分别提取特征 f_s 和 f_g

将提取后的特征连接 $F_c = f_s + f_g$

Step2: 多头注意力模块

for $h \in [H], W_q^h, W_k^h, W_v^h$ 组成

for $h \in [H]$:

$$Q = F_c \times W_q^h$$

$$K = F_c \times W_k^h$$

$$V = F_c \times W_v^h$$

计算每个头注意力输出值:

$$y^h = \text{softmax}\left(\frac{QK^T}{\sqrt{d_K}}\right)V$$

end for

连接所有头部注意力输出值

$$y = [y^1, y^2, \dots, y^H]$$

Step3: PD 诊断识别

$$Dr = \text{MLP}(y)$$

return Dr

end for

2.3 基于余弦混沌的 PD 数据随机拆分和数据编号扰动机制

考虑到用于智能辅助诊断 PD 的语音和步态数据特征包含识别数据主体的大量隐私信息, 为保护数据主体隐私安全, 系统设计过程中融合了一种基于余弦混沌的差分隐私噪声扰动机制。

在语音和步态数据采集完成后, 测试人员将数据上传至系统, 系统接收数据上传指令后, 首先进行两模态数据的处理和噪声扰动, 保证相关数据传输至系统辅助诊断模块的过程安全性, 具体过程如下。

根据帕金森病智能检测所需的目标特征, 对原始数据进行初步处理、标记和矩阵化, 其中语音数据使用声谱矩阵的形式表示, 行表示时间, 列表示频率, 矩阵中每个元素表示相应时间频率下的信号强度, 通过此形式将语音信号转换为数值矩阵; 对于步态数据, 用类似的方式表示, 将每一步的行走数据表示为一个矩阵, 其中行表示不同的特征, 如步长、步幅、频率等, 列表示不同时间点的数据, 通过此形式将步态数据转化为数值矩阵。语音和步态数据形式转换完成后, 通过整合得到两模态特征矩阵 $C = [a, b, \dots, k]$, 若将原始数据直接上传到系统中, 则存在隐私披露风险。对此, 首先将 C 处理成 $(k+2) \times n$ 的矩阵形式, 其中矩阵第一行的所有数字为数据拆分后每列数据的编号, 最后一行为干扰行, 假设干扰初始值为 δ , 结果如公式 (15); 其次引入差分隐私扰动机制, 对第一行数据编号进行加噪处理, 为避免添加随机噪声而导致初始数据无法还原的问题, 使用基于余弦混沌的噪声添加形式, 其中余弦函数值域为 $[-1, 1]$, 为避免不同序号输入值 x 计算出相同的噪声值 f_{noise} , 定义 f_{noise} 的计算公式为:

$$f_{\text{noise}} = \{1 \dots 1\}_n \rightarrow \cos x \quad (14)$$

其中: $x \in (n\pi \sim (n+1)\pi], \{1 \cdots 1\}_n \rightarrow \cos x$ 表示在噪声值 y 前添加 n 个 1, 例如当序号值为 1 时, $1 \in [0, \pi]$, 则 $f_{\text{noise}} = \cos 1 = 0.5403$ 。

$$\begin{bmatrix} a \\ b \\ \vdots \\ k \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & \cdots & n \\ a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_n \\ \vdots & \vdots & \vdots & \vdots \\ k_0 & k_1 & \cdots & k_n \\ \delta & \delta & \delta & \delta \end{bmatrix} \Rightarrow \begin{bmatrix} 0 + \epsilon_0 & 1 + \epsilon_1 & \cdots & n + \epsilon_n \\ a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_n \\ \vdots & \vdots & \vdots & \vdots \\ k_0 & k_1 & \cdots & k_n \\ \delta + \epsilon_0 & \delta + \epsilon_1 & \cdots & \delta + \epsilon_n \end{bmatrix} \Rightarrow \begin{bmatrix} 0 + \epsilon_0 & 1 + \epsilon_1 & 2 + \epsilon_2 & \cdots & n - 1 + \epsilon_{n-1} & n + \epsilon_n \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ b_0 & b_1 & b_2 & \cdots & b_{n-1} & b_n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ k_0 & k_1 & k_2 & \cdots & k_{n-1} & k_n \\ \delta + \epsilon_0 & \delta + \epsilon_1 & \delta + \epsilon_2 & \cdots & \delta + \epsilon_{n-1} & \delta + \epsilon_n \end{bmatrix} \quad (15)$$

在上述基础上, 将随机拆分后的每列数据上传, 系统根据上述加噪方法逆向去掉噪声干扰, 得到恢复后的原始语音和步态数据后进行 PD 的检测和识别。其中随机拆分的数据编号的降噪恢复过程是基于余弦混沌的噪声扰动的逆过程, 见公式 (16)。

$$\begin{bmatrix} 0 + \epsilon_0 & 1 + \epsilon_1 & 2 + \epsilon_2 & \cdots & n - 1 + \epsilon_{n-1} & n + \epsilon_n \\ a_0 & a_1 & a_2 & \cdots & a_{n-1} & a_n \\ b_0 & b_1 & b_2 & \cdots & b_{n-1} & b_n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ k_0 & k_1 & k_2 & \cdots & k_{n-1} & k_n \\ \delta + \epsilon_0 & \delta + \epsilon_1 & \delta + \epsilon_2 & \cdots & \delta + \epsilon_{n-1} & \delta + \epsilon_n \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 1 & \cdots & n \\ a_0 & a_1 & \cdots & a_n \\ b_0 & b_1 & \cdots & b_n \\ \vdots & \vdots & \vdots & \vdots \\ k_0 & k_1 & \cdots & k_n \\ \delta & \delta & \delta & \delta \end{bmatrix} \Rightarrow \begin{bmatrix} a \\ b \\ \vdots \\ k \end{bmatrix} \quad (16)$$

整体过程如下所示:

算法 2: 基于余弦混沌的差分隐私保护算法
 阶段 1: 数据分解加噪过程
 输入: 语音和步态特征矩阵 M , 待传输矩阵数 n
 输出: n 列 PD 两模态数据矩阵

Step1: 数据拆分

$$M \rightarrow \begin{bmatrix} 0 & 1 & \cdots & n \\ m_1 & m_2 & \cdots & m_n \end{bmatrix}$$

Step2: 序号加噪

$$\begin{bmatrix} 0 & 1 & \cdots & n \\ m_1 & m_2 & \cdots & m_n \end{bmatrix} \rightarrow \begin{bmatrix} 0 + \epsilon_1 & 1 + \epsilon_2 & \cdots & n + \epsilon_{n+1} \\ m_1 & m_2 & \cdots & m_n \end{bmatrix}$$

Step3: 矩阵分解

$$\begin{bmatrix} 0 + \epsilon_1 & 1 + \epsilon_2 & \cdots & n + \epsilon_{n+1} \\ m_1 & m_2 & \cdots & m_n \end{bmatrix} \rightarrow \begin{bmatrix} 0 + \epsilon_1 \\ m_1 \end{bmatrix} \cdots \begin{bmatrix} n + \epsilon_{n+1} \\ m_n \end{bmatrix}$$

阶段 2: 数据合并并减噪过程

Step1: 数据矩阵合并

$$\begin{bmatrix} 0 + \epsilon_1 \\ m_1 \end{bmatrix} \cdots \begin{bmatrix} n + \epsilon_{n+1} \\ m_n \end{bmatrix} \rightarrow \begin{bmatrix} 0 + \epsilon_1 & 1 + \epsilon_2 & \cdots & n + \epsilon_{n+1} \\ m_1 & m_2 & \cdots & m_n \end{bmatrix}$$

Step2: 序号降噪

$$\begin{bmatrix} 0 + \epsilon_1 & 1 + \epsilon_2 & \cdots & n + \epsilon_{n+1} \\ m_1 & m_2 & \cdots & m_n \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & \cdots & n \\ m_1 & m_2 & \cdots & m_n \end{bmatrix}$$

Step3: 数据恢复

$$\begin{bmatrix} 0 & 1 & \cdots & n \\ m_1 & m_2 & \cdots & m_n \end{bmatrix} \rightarrow M$$

3 实验结果与分析

系统设计完成后, 为了验证和优化基于多头注意力机制的 PD 智能辅助诊断模型性能, 进一步进行了测试和验证。

3.1 数据集

实验中使用来自 mPower 研究中的两模态语音和步态数据集。该数据集包括 65 022 个独特的任务, 包括 5 826 个个体受试者, 其中每条数据中均包含 10 秒的语音样本。步态数据集存储为 JavaScript 对象表示法 (JSON) 文件。在本实验中, PD 智能检测模型的输入数据为处理和融合后的语音和步态特征数据。

3.2 实验环境设置

本实验在浪潮服务器中运行, 使用的库和编程语言分别为 Pytorch1. 10. 1 和 Python3. 7. 0。实验的硬件环境为 64 位 Intel (R) Xeon (R) Silver 4210R CPU@2.40 GHz 处理器和 32 GB RAM 模拟环境来训练和测试 MHA-CNN。在模型训练中, 我们将实验数据按照 8 : 1 : 1 的比例分为训练集、验证集和测试集。

3.3 MHA-CNN 性能评估指标

在本节中, 分别使用准确率、 F_1 -score、精确度和召回率^[21-22]作为模型性能的评估指标。其中模型精度和召回率的计算公式为:

$$Precision = \frac{TP}{TP + FP} \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

模型准确率的计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

其中: TP 表示 PD 被正确识别的样本数量, FP 表示被误报的非 PD 样本数量, TN 表示非 PD 被正确识别的样本数量, FN 为被漏报的 PD 样本数量。

$$F_1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (20)$$

其中: F_1 -score 是精度和召回率的加权求和平均值, 精确度表示所有 PD 阳性样本中被准确预测为阳性样本的百分

比, 召回率表示所有 PD 阳性样本中被正确预测为阳性样本的检出率。

3.4 实验分析

为测试和验证 MHA-CNN 模型的性能, 本节进行了两模态消融实验和模型性能对比实验, 分别从精确度、准确率、召回率、损失值、 F_1 -score 等 5 个评估指标系统评估了 MHA-CNN 性能, 具体实验结果如下。

3.4.1 两模态消融实验

图 3 是 MHA-CNN 在训练集中的训练结果。如图 3 (a~b) 分别为模型 AP 曲线图和 ROC 曲线图, 通过对比, 添加多头注意力机制后的 MHA-CNN 模型 AP 曲线与 ROC 曲线下面积均大于未添加注意力的曲线, 表明 MHA-CNN 模型性能更好; 如图 3 (c), 随着迭代轮次的增加, 添加注意力的模型准确率快速上升, 并趋于稳定, 其中峰值准确率约为 0.99; 如图 3 (d) 所示, 随着模型训练轮数的增加, 添加注意力的模型损失值下降速度更快, 并快速收敛, 损失值约为 0.32, 性能均优于未添加注意力机制的模型。

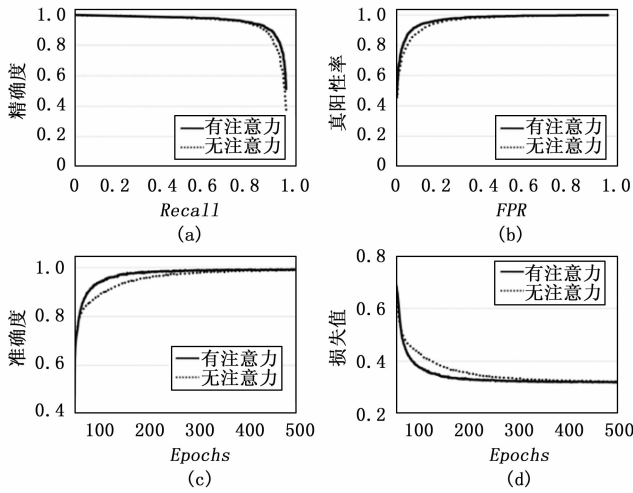


图 3 MHA-CNN 两模态消融实验结果

3.4.2 模型性能对比实验

为进一步验证 MHA-CNN 模型性能, 研究中对 MHA-CNN 模型与传统的 LWF 模型的性能进行了对比, 对比实验结果如图 4 所示。

图 4 (a~d) 显示了 MHA-CNN 和 LWF 在测试集中的运行结果。MHA-CNN 的 Accuracy、Precision、Recall 和 F_1 -score 分别为 0.913、0.908、0.904 和 0.906, 要明显高于 LWF 的 0.643、0.50、0.321 和 0.391, 在基于两模态数据的 PD 智能辅助诊断效果层面要明显优于 LWF 模型。图 3 (e) 表明, 随着测试次数的增加, MHA-CNN 的准确率快速上升并收敛。实验结果表明, 基于多头注意力机制的帕金森病多模态远程检测系统能够满足 PD 大规模早期安全筛查要求。

4 结束语

本文设计了一种基于多头注意力机制的帕金森病多模

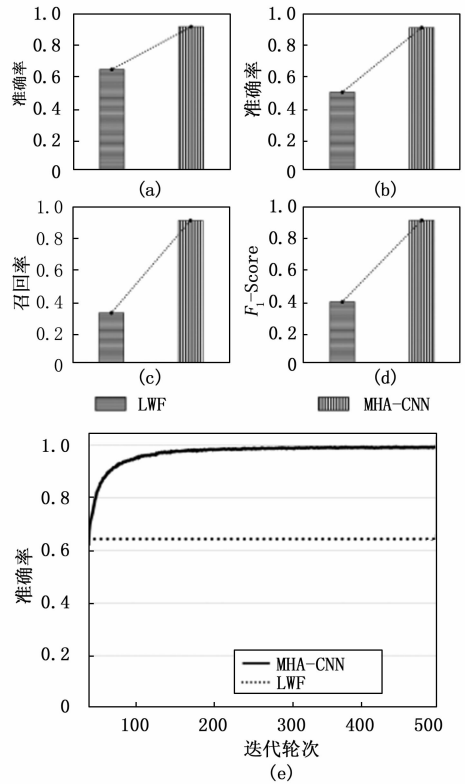


图 4 模型性能对比实验结果

态安全远程辅助检测系统。研究中通过在卷积神经网络 CNN 后融合多头注意力机制, 提高了卷积神经网络算法模型的多模态数据特征提取、融合和识别能力。同时考虑到 PD 患者语音和步态数据特征所包含的能够唯一识别数据主体的隐私信息, 研究中使用了一种基于余弦混沌的差分隐私保护噪声扰动方式, 在数据传输前将语音和步态数据随机拆分并编号, 通过向编号中添加噪声的形式, 保证数据传输过程的安全性。为了验证 MHA-CNN 模型的性能, 本文进行了两模态消融实验和对比实验, 仿真实验结果表明, MHA-CNN 的准确率、精度等高于 0.9, 且模型的准确率和损失随着训练和测试的轮次增加, 均快速收敛并趋于稳定。实验结果达到了 PD 检测识别的预期目标, 在提高 PD 远程诊疗准确性和稳定性的同时, 保证了 PD 数据的隐私安全性。

在后续研究中, 将进一步融合用于 PD 临床检验的文本数据等, 开展更高模态的 PD 智能辅助检测研究, 持续提高 PD 早期检测准确率。同时, 开展 PD 辅助检测过程的安全性研究, 确保 PD 多模态远程检测过程的安全性, 不断提升 PD 辅助诊疗结果对于 PD 临床诊疗的决策支持作用和价值。

参考文献:

[1] KLEMANN C, MARTENS G, SHARMA M, et al. Integrated molecular landscape of Parkinson's disease [J]. npj Parkinson's Disease, 2017, 3 (1): 14.

[2] MARINI K, SEPPI K, TSCHIDERER L, et al. Application of

- the upated movement disorder society criteria for prodromal parkinson's disease to a population based 10th ear study [J]. *Movement Disorders*, 2021, 36 (6): 1464-1466.
- [3] 陈真真, 张玉虎. 帕金森病的临床亚型研究进展 [J]. *中华神经科杂志*, 2022, 55 (3): 248-253.
- [4] 谢俊泉, 赵 欢, 曹军义, 等. 帕金森病步态的可穿戴多源数据采集与特征识别方法 [J]. *西安交通大学学报*, 2023, 57 (6): 29-38.
- [5] 许之珩, 张梦翰, 王 坚. 语音声学分析在帕金森病中的诊断价值 [J]. *四川大学学报 (医学版)*, 2022, 53 (4): 726-731.
- [6] ALMEIDA J S, REBOUCAS F, PEDRO P, et al. Detecting parkinson's disease with sustained phonation and speech signals using machine learning techniques [J]. *Pattern Recognition Letters*, 2019, 125: 55-62.
- [7] TSANAS A, LITTLE MA, MCSHARRY PE, et al. Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease [J]. *IEEE Transactions on Biomedical Engineering*, 2012, 59 (5): 1264-1271.
- [8] WROGE T J, ÖZKANCA Y, DEMIROGLU C, et al. Parkinson's disease diagnosis using machine learning and voice [J]. *IEEE Signal Processing in Medicine and Biology Symposium*, 2018: 1-7.
- [9] 朱家英, 徐志京. 基于动态注意力机制和多模态循环融合的帕金森氏症识别 [J]. *计算机应用研究*, 2023, 40 (2): 481-487.
- [10] YOO I C, LEE K, LEEM S, et al. Speaker anonymization for personal information protection using voice conversion techniques [J]. *IEEE Access*, 2020, 8: 198637-198645.
- [11] 顾理平. 智能生物识别技术: 从身份识别到身体操控——公民隐私保护的视角 [J]. *上海师范大学学报 (哲学社会科学版)*, 2021, 50 (5): 5-13.
- [12] WANG X, LI F, XIE Y, et al. Gait and respiration-based user identification using Wi-Fi signal [J]. *IEEE Internet of Things Journal*, 2022, 9 (5): 3509-3521.
- [13] 蒋 伟, 王瑞锦, 余苏喆, 等. 基于步态识别的移动设备身份认证模型 [J]. *电子科技大学学报*, 2019, 48 (2): 272-277.
- [14] HO N H, YANG H J, KIM S H, et al. Multimodal approach of speech emotion recognition using multi-level multi-head fusion attention-based recurrent neural network [J]. *IEEE Access*, 2020, 8: 61672-61686.
- [15] 于重重, 吴佳佳, 陈运兵, 等. 基于多头注意力机制的端到端土家语语音识别 [J]. *计算机仿真*, 2022, 39 (3): 258-262, 282.
- [16] CYNTHIA D. Differential privacy in proceedings of the 33rd international conference on automata [C] // *Languages and Programming-Volume Part II (ICALP'06)*. Springer-Verlag, Berlin, Heidelberg, 2006, 1-12.
- [17] CHANGQIN Q, KANG R, ZHWEI L, et al. End-to-end deep learning approach for parkinson's disease detection from speech signals [J]. *Biocybernetics and Biomedical Engineering*, 2022, 42: 556-574.
- [18] LIU Y, REDDY M K, et al. Automatic assessment of parkinson's disease using speech representations of phonation and articulation [J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2023, 31: 242-255.
- [19] SIVARANJINI S, SUJATHA C M. Deep learning based diagnosis of parkinson's disease using convolutional neural network [J]. *Multimedia Tools Appl*, 2020, 79: 15467-15479.
- [20] SAN-SEGUNDO R, NAVARRO-HELLIN H, Torres- Sánchez R, et al. Increasing robustness in the detection of freezing of gait in parkinson's disease [J]. *Electronics*, 2019, 8 (2): 119.
- [21] 莫桂江, 颜增显. 基于 OBS-CNN 的视觉 SLAM 闭环检测方法研究 [J]. *火力与指挥控制*, 2022, 47 (12): 93-97.
- [22] 李建威, 吕晓琪, 谷 宇. 基于改进 ConvNeXt 的皮肤镜图像分类方法 [J]. *计算机工程*, 2023, 49 (10): 239-246.
- [8] 朱文杰, 凌 鹤, 杨守鹏. 基于双目视觉的车身焊点定位误差补偿研究 [J]. *应用光学*, 2021, 42 (1): 79-85.
- [9] 黄知超, 梁国丽, 朱芳来. 微芯片焊点快速识别与定位方法 [J]. *半导体技术*, 2020, 45 (3): 236-243.
- [10] 刘 涛, 左继业, 张晓森, 等. 环形焊缝的双目视觉定位与跟踪研究 [J]. *机械科学与技术*, 2021, 40 (6): 928-933.
- [11] 田林琳. 基于深度学习及 GPU 计算的航天器故障检测技术 [J]. *计算机测量与控制*, 2020, 28 (5): 1-4.
- [12] 刘哲成, 郭丽娟. 基于分层滤波算法的无人机控制系统故障检测技术 [J]. *计算机测量与控制*, 2020, 28 (5): 23-26.
- [13] 朱道俊, 徐湛楠, 马婷婷, 等. 基于视觉和轮速计紧耦合的轮式机器人定位算法研究 [J]. *传感技术学报*, 2021, 34 (7): 896-903.
- [14] 高海伟. 基于本征正交分解法的焊接模型位移重构 [J]. *组合机床与自动化加工技术*, 2020 (1): 22-24.
- [15] 晁佳佳. 焦化厂自动化仪表的故障预防及其维护策略分析 [J]. *中国科技期刊数据库工业 A*, 2023 (6): 160-162.
- [16] 杜康宁, 宁少慧, 邓功也. 基于视觉 Transformer 的滚动轴承智能故障诊断 [J]. *组合机床与自动化加工技术*, 2023 (4): 96-99.
- [17] 李亚伟, 窦 明, 康来坡. 煤矿机电设备管理中机械故障检测诊断技术的应用 [J]. *内蒙古煤炭经济*, 2023 (2): 163-165.
- [18] 赵嗣芳, 宋 强, 张艳明, 等. 基于改进 DFA 和 LDA 的永磁同步电机机械故障检测 [J]. *北京理工大学学报*, 2023, 43 (1): 61-69.
- [19] 田 园, 黄祖源, 张 航, 等. 基于混合特征和集成极限学习机的高压断路器机械故障检测 [J]. *电子器件*, 2023, 46 (1): 203-209.
- [20] 李 欢, 吕 勇, 袁 锐, 等. 基于深度卷积神经网络的滚动轴承迁移故障诊断 [J]. *组合机床与自动化加工技术*, 2023 (2): 90-94.