

基于噪声混淆增强特征鲁棒性的脑疾病预测

郝小可, 谭麒豪, 李家旺, 郭迎春, 于明

(河北工业大学人工智能与数据科学学院, 天津 300401)

摘要: 随着医学影像数据的不断发展, 纵向数据分析逐渐成为了解和跟踪阿尔茨海默病(Alzheimer's disease, AD)发病过程的重要研究方向。目前已经提出了许多纵向数据分析方法, 其中多任务学习得到广泛应用, 它能够集成多个时间点的影像数据, 提高模型的泛化能力。大多数现有的方法能够识别不同时间点的共享特征, 但这些特征中会包含一定的噪声。与此同时, 不同时间点进展的潜在关联仍未得到充分的探索。本文提出了一种基于参数分解和关系诱导的多任务学习(Parameter decomposition and relation-induced multi-task learning, PDRIMTL)方法, 以此从纵向数据中识别特征。该方法不仅能够识别去除噪声后的共享特征, 提高共享特征的鲁棒性, 而且能够对不同时间点的内在关联进行建模。结果表明, 在不同时间点的结构磁共振成像(Structural magnetic resonance imaging, sMRI)数据上, 该模型能够有效提高对AD鉴别的准确性。

关键词: 阿尔茨海默病; 多任务学习; 参数分解; 关系诱导正则化; 纵向数据

中图分类号: TP391 **文献标志码:** A

Brain Disease Prediction Based on Noise Confusion to Enhance Robustness of Features

HAO Xiaoke, TAN Qihao, LI Jiawang, GUO Yingchun, YU Ming

(School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China)

Abstract: With the continuous development of medical imaging data, longitudinal data analysis is gradually becoming an important research direction to understand and trace the process of the Alzheimer's disease (AD). At present, many longitudinal data analysis methods have been proposed, among which multi-task learning is widely used, which can integrate imaging data of multiple time points to improve the generalization ability of the model. Most existing methods can identify shared features at different time points, but these features will contain a certain amount of noise. At the same time, potential associations of disease progression at different time points remain under explored. In this paper, we propose a parameter decomposition and relation-induced multi-task learning (PDRIMTL) method to identify features from longitudinal data. The method can not only identify shared features after noise removal and improve the robustness of shared features, but also can model the intrinsic associations between different time points. The results show that the model can effectively improve the accuracy of AD identification on structural magnetic resonance imaging (sMRI) data at different time points.

基金项目: 国家自然科学基金(61806071); 河北省自然科学基金(F2020202025)。

收稿日期: 2022-05-10; **修订日期:** 2022-07-13

Key words: Alzheimer's disease; multi-task learning; parameter decomposition; relation-induced regularization; longitudinal data

引言

阿尔茨海默病(Alzheimer's disease, AD)是最常见的痴呆形式,是一种隐匿的进行性神经退行性疾病,主要影响人体大脑的认知和记忆功能,甚至可能导致死亡^[1]。AD不仅给医疗保健系统带来了巨大的经济压力,而且使得患者及其家人经历了巨大的心理和情感压力。目前还没有能够逆转AD进展的治疗方法,但及时诊断和治疗AD对于延缓疾病进展仍具有重要的意义^[2]。

近几十年来,基于结构磁共振成像(Structural magnetic resonance imaging, sMRI)的纵向数据被广泛用于预测认知状态,纵向数据研究通常采用患者在不同时间点的检测数据。目前许多认知测量方法被设计用来评估认知衰退,例如,简易精神状态评分量表(Mini-mental state examination, MMSE)和阿尔茨海默病评估量表认知亚量表(Alzheimer's disease assessment scale cognitive subscale, ADAS-Cog),它们可以用来揭示AD的进展。据报道,MMSE与潜在的AD病理和神经退行性机制相关^[3],而ADAS-Cog是AD认知功能评估试验的金标准^[4]。

目前已经存在一些专门用于AD疾病进展建模的特征选择工作,这些工作已被证实是反映多个时间点预测任务之间关联的有效方法。这些方法一般可分为两类:单任务学习和多任务学习(Multi-task learning, MTL)。单任务学习通常分别估计不同时间点的AD疾病进展情况。也就是说,在单任务学习中,既没有探索成像特征与时间点之间的关联,也没有探索不同预测时间点之间的相关性。例如,Duchesne等^[5]提出使用线性回归模型从基线sMRI数据中预测一年的MMSE变化,并发现基线数据与一年的认知变化高度相关;Wang等^[6]开发了一个线性相关的向量回归模型来预测连续的临床变量。由于单任务学习将每个时间点视为一个独立的任务,不考虑不同预测时间点之间的内在关联,因此它在预测脑疾病进展方面的表现并不理想。

作为另一个努力方向,MTL^[7]不仅可以发现不同预测时间点之间的潜在关系,而且还可以提高模型的泛化能力。MTL已被广泛用于脑疾病分析领域的认知状态预测^[8-11]。基于稀疏诱导规范的多任务特征学习(Multi-task feature learning, MTFL)能够选择来自sMRI特征的鉴别特征子集,摒弃一些对于模型学习没用的特征。例如,Zhou等^[8]提出将时间平滑性整合到MTL模型中,用于预测疾病进展和选择疾病进展的生物标志物;Jie等^[11]开发了一种用于纵向数据分析的时间约束的组稀疏学习方法,反映了相邻时间点的数据之间的平滑变化;Nie等^[10]设计了一种新的多源多任务学习方法,利用源一致性和时间平滑性的先验知识来评估疾病状态,并验证了其对AD患者进展预测的有效性。现有的方法就是将一些有效的框架进行应用,如Cao等^[12]提出了一种具有 $L_{2,1}$ 范数和 $G_{2,1}$ 范数的稀疏多任务学习方法,考虑了多个神经成像测量之间的协方差结构,可以识别与AD认知相关的脑区域;Jiang等^[13]考虑了相关感知的稀疏和低秩约束正则化,能够准确预测患者在不同时间点的认知得分。

现有的MTL方法通过学习多个相关时间点的影像数据,并引入特定的正则化项来建模。在预测认知分数时,通常希望融合不同模态或不同时间点的影像数据,以此发现不同时间点之间的共性。因此 $L_{2,1}$ 范数和 $G_{2,1}$ 范数在纵向数据分析中广泛使用,对参数矩阵施加结构稀疏^[10-11],能够发现不同时间点之间的共享特征。然而,大多数现有的模型在建模疾病进展时没有考虑到相邻预测时间点对于疾病发展的必然联系,并且现有方法在集中于对不同时间点共享特征的探索时,忽略了其中包含的噪声对预测性能的影响。

基于以上分析,本文提出了一种基于参数分解和关系诱导的多任务学习方法。该方法通过参数分

解计算不同参数相关权值,以此识别多个时间点中包含的噪声和共享特征,将摒弃噪声后的共享特征用于预测能够更好地提高性能。此外,为了自动学习不同预测时间点之间的关系,本文在模型中使用关系诱导^[14]。吸收了多任务学习、参数分解以及关系诱导的优点,MTL框架能够集成多个时间节点的sMRI数据,参数分解对不同的参数施加不同的正则化项,这是非常有意义的,而关系诱导能够引导关系信息的学习。

本文方法主要有以下3点贡献:

(1) MTL框架:能够集成多个时间点的影像数据,提高模型的泛化能力。

(2) 参数分解:可以对不同的参数矩阵施加不同的正则化项,以不同的方式惩罚不同时间点的噪声和共享特征。

(3) 关系诱导:能够有效地从数据中引入关系信息,以此来指导知识传递,能够发现不同预测时间点之间的潜在联系。

1 相关方法

1.1 多任务学习模型

在纵向数据研究中,MTL模型常用于结合不同时间点的影像数据,能够发现不同时间点之间的潜在关联,提高模型的泛化能力。多任务学习模型可表示为

$$\min_Z \sum_{i=1}^T \|X_i W_i - Y_i\|_2^2 + \alpha \|W\|_{2,1} \quad (1)$$

式中: $X_i \in \mathbf{R}^{N \times d}$,假设训练对象的数量为 N ,每个受试者在 T 个不同时间有对应的 d 维成像数据; $Y_i \in \mathbf{R}^{N \times 1}$,表示受试者在 T 个不同时间点对应的评分量表。基于 $L_{2,1}$ 范数的多任务学习是同时解决回归和稀疏性问题的有效方法之一,在纵向数据研究中, $L_{2,1}$ 范数常用于寻找不同任务之间最具鉴别性的特征; $W \in \mathbf{R}^{d \times T}$,式中 $\|W\|_{2,1}$ 定义为 $\sum_i \sqrt{\sum_j w_{ij}^2} = \sum_i \|w^i\|_2$, L_1 范数保证了模型的稀疏性, L_2 范数保证了所有时间点之间的相似性模式,确保相同的脑区被选中。

1.2 基于参数分解的多任务学习模型

在MTL模型的基础上,本文对式(1)中的标准权重 W 进行参数分解,即能够去除不同时间点的噪声影响,从而进一步提高模型对鲁棒特征的分析能力。模型可以表述为

$$\min_{Z, V} \sum_{i=1}^T \|X_i(Z_i + V_i) - Y_i\|_2^2 + \alpha \|Z\|_{2,1} + \beta \|V\|_{1,1} \quad (2)$$

在模型中,标准权重 W 被分解为两部分,即 $W = Z + V$,其中 Z 是不同时间点的共享权重, V 是噪声权重,式中 α, β 为非负参数,式中 $Z \in \mathbf{R}^{d \times T}$, $V \in \mathbf{R}^{d \times T}$ 。得益于参数分解,可以针对不同的组成部分施加不同的惩罚。

本文定义了公式(3),这有助于保证不同时间点的相似模式,能够选择不同时间点之间一致性的共享特征。

$$\|Z\|_{2,1} = \sum_{i=1}^d \|z^i\|_2 = \sum_{i=1}^d \sqrt{\sum_{t=1}^T (z_t^i)^2} \quad (3)$$

本文定义了公式(4),在识别不同时间点噪声的同时,能够对噪声进行稀疏。

$$\|V\|_{1,1} = \sum_{i=1}^d \sum_{t=1}^T |v_t^i| \quad (4)$$

1.3 基于参数分解和关系诱导的多任务学习模型

到目前为止,基于参数分解的模型已经能够识别噪声和不同时间点一致性的共享特征。此外,为了自动学习不同预测任务之间的关系,本文在提出的模型中使用关系诱导项^[14],定义如下

$$\begin{aligned} & \text{tr}(W\Omega^{-1}W^T) \\ \text{s.t. } & \Omega \geq 0, \text{tr}(\Omega) = 1 \end{aligned} \tag{5}$$

式中: $\text{tr}(\cdot)$ 表示方阵的迹, Ω^{-1} 表示矩阵的逆,定义为任务协方差矩阵,在纵向数据分析中诱导正确的关系,这有利于不同权重的学习。不同时间点临床评分之间的相关性有助于反映来自相邻时间点数据之间的平滑变化。用 $\Omega \geq 0$ 来将 Ω 限制为半正定矩阵,用 $\text{tr}(\Omega) = 1$ 惩罚 Ω 的复杂度。

通过合并在等式中定义的关系诱导项(式(5)和式(2)),目标函数可表示为

$$\begin{aligned} \min_{Z, V, \Omega} & \sum_{t=1}^T \|X_t(Z_t + V_t) - Y_t\|_2^2 + \alpha \|Z\|_{2,1} + \beta \|V\|_{1,1} + \gamma \text{tr}(Z\Omega^{-1}Z^T) \\ \text{s.t. } & \Omega \geq 0, \text{tr}(\Omega) = 1 \end{aligned} \tag{6}$$

式中: α, β, γ 为正则化参数,以控制参数分解和关系诱导正则化之间的平衡; Ω 由不同时间点的共享权重 Z 来引导。在下一节中,将开发一个有效的优化算法来求解在等式中定义的目标函数。

1.4 模型框架

基于参数分解和关系诱导的多任务学习框架如图1所示,采用受试者在BL、M06、M12、M24这4个时间点的sMRI数据作为输入,4个时间点对应的认知评分(MMSE、ADAS-Cog)作为输出。模型中优化的参数为 Z, V, Ω ,参数分解将权重 W 分解为 V 和 Z ,其中 Z 代表不同时间点的共享权重, V 代表噪声权重, Ω 代表不同预测时间点之间的相关性权重,它由共享权重 Z 引导。

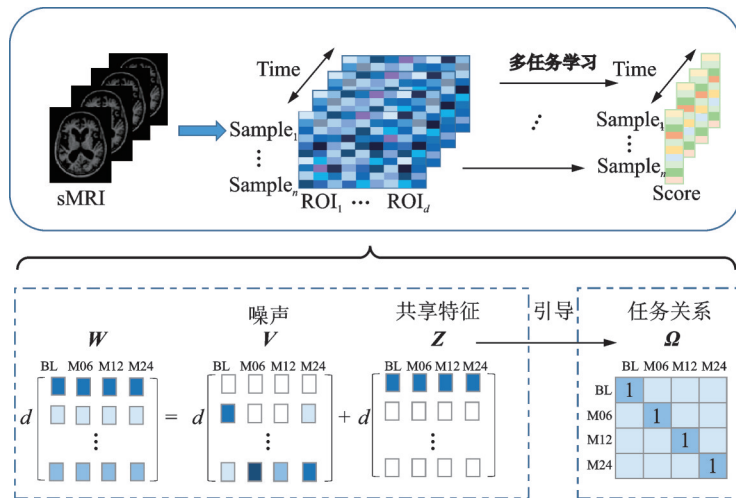


图1 基于参数分解和关系诱导的多任务学习框架

Fig.1 Multi-task learning framework based on parameter decomposition and relation induction

2 算法求解

本文的目标是求解噪声权重 V 、共享权重 Z 以及不同预测时间点的相关性权重 Ω 。目标函数(式6)是一个凸优化问题,建议使用交替算法来求解。首先用固定值 V, Ω 优化 Z ,然后用固定值 Z, Ω 优化 V ,同理优化 Ω 。

更新 Z :

如果 V 和 Ω 固定为常数,式(6)可以简化为

$$\min_Z \sum_{t=1}^T \|X_t Z_t - Y_t\|_2^2 + \alpha \|Z\|_{2,1} + \gamma \text{tr}(Z \Omega^{-1} Z^T) \quad (7)$$

首先在式(7)中展开二次项

$$\min_Z \sum_{t=1}^T (\|X_t\|_2^2 - 2Z_t^T X_t^T Y_t + \|Y_t\|_2^2) + \alpha \|Z\|_{2,1} + \gamma \text{tr}(Z \Omega^{-1} Z^T) \quad (8)$$

因此,很容易得到了式(7)的上界,即

$$(X_t^T X_t + \alpha D + \gamma \Omega^{-1}) z_t = X_t^T Y_t \quad (9)$$

那么可以得到 $z_t = (X_t^T X_t + \alpha D + \gamma \Omega^{-1})^{-1} X_t^T Y_t$, D 是一个对角矩阵,可以表示为 $\frac{1}{2|z_t^i|}$ ($i=1, \dots, d$)。

同理可以得到

$$(X_t^T X_t + \beta Q) v_t = X_t^T Y_t \quad (10)$$

那么可以得到 $v_t = (X_t^T X_t + \beta Q)^{-1} X_t^T Y_t$, Q 是一个对角矩阵,可以表示为 $\frac{1}{2|v_t^i|}$ ($i=1, \dots, d$)。

当 Z 和 V 固定时, Ω 的优化问题可以表述为

$$\min_{\Omega} \text{tr}(Z \Omega^{-1} Z^T) \quad \text{s.t. } \Omega \geq 0, \text{tr}(\Omega) = 1 \quad (11)$$

根据文献[13,14],通过取式(11)的偏导数,可以得到 Ω 的闭形解为

$$\Omega = \frac{(Z^T Z)^{\frac{1}{2}}}{\text{tr}(Z^T Z)^{\frac{1}{2}}} \quad (12)$$

算法 基于参数分解和关系诱导的多任务学习算法流程

输入:sMRI数据 $X_t \in \mathbf{R}^{n \times d}$, 认知评分 $Y_t \in \mathbf{R}^{n \times 1}$, $t=1, \dots, T$, 正则化参数 α, β, γ 。

输出:不同时间点的共享权重 Z , 噪声权重 V , 不同预测时间点之间的相关性权重 Ω 。

- (1) 初始化权重 Z, V 和 Ω ;
- (2) while 不收敛 do
- (3) 更新 Z ;
- (4) 通过式(9)求解 $Z, Z \in \mathbf{R}^{d \times T}$;
- (5) 更新 V ;
- (6) 通过式(10)求解 $V, V \in \mathbf{R}^{d \times T}$;
- (7) 更新 Ω ;
- (8) 通过式(11)求解 $\Omega, \Omega \in \mathbf{R}^{T \times T}$;
- (9) end while

3 实验分析与结果

为了评估所提模型的有效性,本文使用ADNI数据集中550个受试者在4个不同的时间点(即基线、6个月、12个月和24个月)的sMRI数据和相应的认知评分(即MMSE和ADAS-Cog)进行实验。实验中,分别预测了MMSE和ADAS-Cog评分。受试者有3种诊断类别,即健康对照(Healthy control, HC)、轻度认知障碍(Mild cognitive impairment, MCI)和AD。

将本文方法与一些现有的主流方法进行比较,其中包括Lasso^[15]、MTFL^[16]、多任务关系学习(Multi-task relationship learning, MTRL)^[14, 17]和参数分解的多任务特征学习(Parametric decomposition

multi-task feature learning, PDMTFL)^[18-19]等方法。整个数据集划分为10个大小大致相同的子集,选择1个子集进行测试,剩下的9个子集进行训练,根据测试数据的结果评估方法性能。

使用支持向量回归器(Support vector regression, SVR)默认参数(即 $C=1$)^[5]进行回归预测。为了选择最优参数,使用5折交叉验证,使用 $\{10^{-5}, 10^{-4}, \dots, 10^5\}$ 来选择参数。经过交叉验证后,选择在训练数据上性能最好的参数进行测试。为了避免数据集分区时的随机性,重复该过程10次,以展示重复交叉验证结果平均后的性能。

3.1 对比方法

(1) Lasso:通过 L_1 范数进行正则化,进行变量选择以提高预测和回归性能。它通过对不同时间点的认知评分分别建模进行回归预测,是估计未来疾病状态最广泛的方法之一。

(2) MTFL:对权重矩阵施加 $L_{2,1}$ 范数,能够从多个任务中选择共享特征,它鼓励多个任务共享相似的稀疏模式。

(3) MTRL:在正则化框架下学习任务关系矩阵,它是对单任务学习正则化框架的一种新的推广,不同时间点的关系矩阵和模型参数可以进行迭代学习。

(4) PDMTFL:针对不同的权重施加不同的惩罚项,鼓励不同的稀疏结构,这是非常有意义的。

3.2 评估参数

对回归模型进行性能评估时,通常采用均方根误差(Root mean squared error, rMSE)、归一化均方根误差(Normalized mean squared error, nMSE)指标进行评判。nMSE常应用于MTL研究中^[20],即均方误差除以真实值的方差。rMSE广泛用于测量预测和实际临床评分之间的回归和关联分析性能^[21]。对于nMSE和rMSE而言,值越小,性能越好。本文实验采用rMSE、nMSE以及不同预测时间点的相关性矩阵 Ω 作为评价指标。

$$\text{rMSE} = \sqrt{\frac{\|Y - \hat{Y}\|_2^2}{n}} \quad (13)$$

$$\text{nMSE} = \frac{\sum_{i=1}^T \|Y_i - \hat{Y}_i\|_2^2 / \sigma(Y_i)^2}{\sum_{i=1}^T n_i} \quad (14)$$

不同方法的对比实验结果如表1所示,其中最好的结果用粗体显示。

表1 不同方法实验对比
Table 1 Experimental comparison of different methods

评估指标	Lasso	MTFL	PDMTFL	MTRL	PDRIMTL	
nMSE	0.685 2	0.643 0	0.621 3	0.601 8	0.586 4	
Baseline rMSE	2.206 7	2.091 0	1.912 4	1.891 3	1.773 0	
MMSE	M06 rMSE	2.806 8	2.683 1	2.456 4	2.431 2	2.271 5
	M12 rMSE	3.307 4	3.130 2	2.967 3	2.901 5	2.281 9
	M24 rMSE	4.023 1	3.957 1	3.794 1	3.721 8	3.651 2
ADAS-Cog	nMSE	0.628 5	0.605 9	0.589 6	0.591 3	0.537 5
	Baseline rMSE	7.483 6	7.361 7	7.133 0	7.078 7	6.964 5
	M06 rMSE	8.013 7	7.872 2	7.758 1	7.722 3	7.657 2
	M12 rMSE	8.541 0	8.438 1	8.273 1	8.112 4	7.972 1
	M24 rMSE	10.104 4	9.724 1	9.517 3	9.455 1	9.321 4

结果表明,本文方法在nMSE和rMSE评判标准上均优于其他比较方法,证明了该方法在联合预测认知评分方面的有效性。对于所有方法,都采用10折交叉验证来估计预测性能。由于参数值过小会导致处罚不足,参数值过大会导致处罚过度,所以调整 α 、 β 、 γ 范围 $\{10^{-5}, 10^{-4}, \dots, 10^5\}$ 。通过表1中评价指标,可以得出以下观察结果:首先,与单任务学习方法Lasso相比,使用多个时间点进展信息的学习方法明显都能取得更好的表现。这说明,多个时间点的进展信息有利于预测疾病的进展,同时多个时间点的影像数据包含更多的疾病渐进信息。其次,本文方法能够自动建模不同预测时间点的内在关系,在nMSE、rMSE以及不同预测时间点的相关性方面始终优于4种竞争方法。本文方法在估计MMSE和ADAS-Cog分数时的平均(即4个时间点)nMSE分别为0.537 5和0.586 4,这些结果是有效的。对表1实验结果进行了可视化,结果如图2和3所示,其中Baseline表示基线,M06、M12和M24分别表示基线后6、12和24个月。从图2和3中可更为直观地看出,本文方法的性能均优于其他方法。

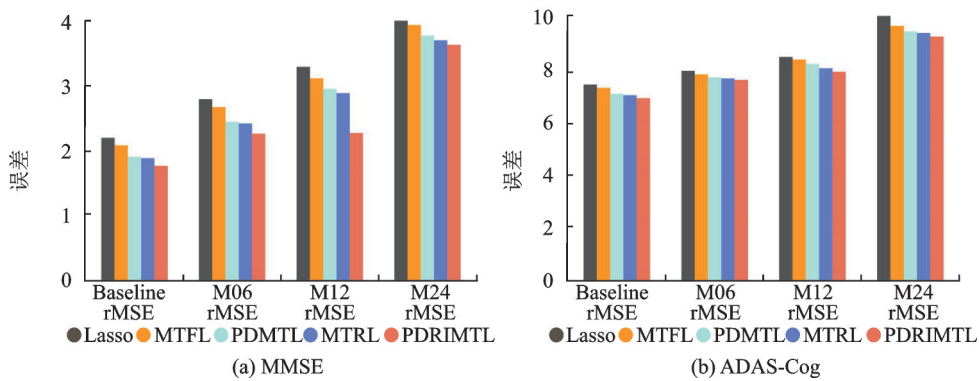


图2 不同方法在rMSE下的比较

Fig.2 Comparison of different methods under rMSE

3.3 学习任务关系

实验估计了MTRL和本文方法分别在MMSE和ADAS-Cog中学到的任务关系,结果如图4、5所示,它们是对 Ω 的可视化。需要注意的是,每个回归任务都对应于一个时间点,BL表示基线;M06、M12和M24分别表示基线后6、12和24个月。其中颜色条中的黄色表示一个较高的相关系数。从图中可以看出回归任务在4个时间点的相关性存在显著差异,采用不同的评分量表,MTRL在不进行特征选择的情况下,4个时间点学习到的关系都较低,这与多个时间点往往发挥相似的作用相悖。通常认为

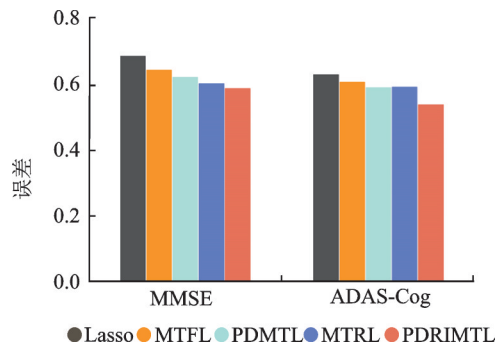


图3 不同方法在nMSE下的比较

Fig.3 Comparison of different methods under nMSE

为多个时间点往往扮演相似的角色,它们的相关性应该是较高的。本文方法在通过等式中的正则化项,在摒弃不同时间点存在的噪声特征之后,识别到的不同时间点的共享特征能够更好地提高个体回归变量的学习表现,学习到不同时间点之间的真实关系。实验结果中,也可以发现相邻时间点的相关

性与样本数量有很大的关系,相邻时间点中样本数量多的相关性通常是较高的。

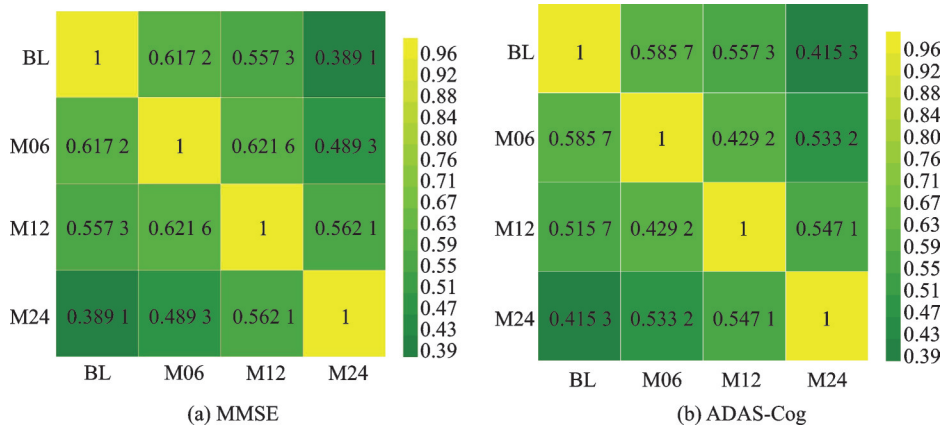


图4 MTRL方法学习的4个回归任务之间的相关矩阵

Fig.4 Correlation matrix among four regression tasks learned by MTRL method

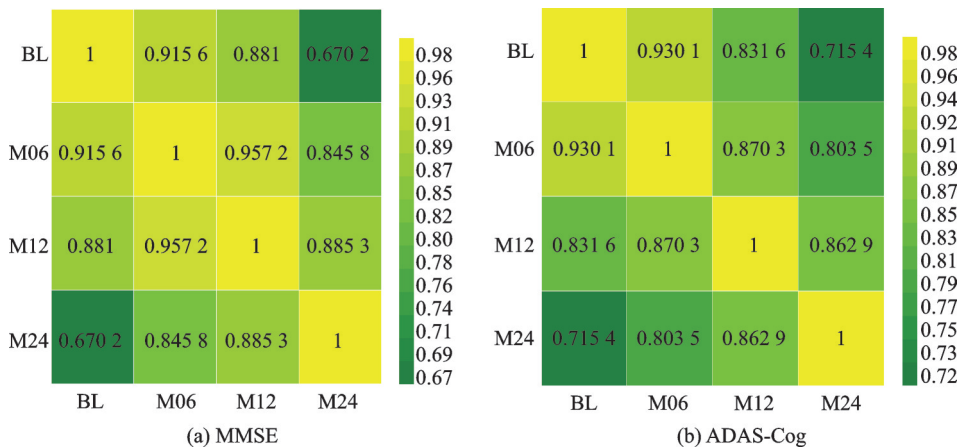


图5 本文方法学习的4个回归任务之间的相关矩阵

Fig.5 Correlation matrix among four regression tasks learned by the proposed method

3.4 判别脑区分析

识别与AD进展高度相关的生物标志物子集非常有意义。因此,展示了由本文方法所确定的前10个大脑区域。值得注意的是,由于在每个10折交叉验证中被选择的大脑区域是不同的,所以选择累积绝对权重作为大脑区域在回归任务中的贡献的指标。图6为所提方法识别的前10个重要脑区,所选出的脑区名称如表2所示。从表2可以看出,两个回归任务所选择的脑区都包括信息最丰富的大脑区域海马体和杏仁核,相同区域的脑区被选中,这也证实了本文方法识别的一致性。通过AD建模和测量认知结果,证实了两者在识别大脑状况方面发挥着重要作用,与阿尔茨海默病高度相关。海马体^[22]与陈述性记忆密切相关,而杏仁核^[23]在记忆情感重大经历中起着重要作用。选定的脑区中颞叶和颞中区^[8]也对AD诊断敏感,在以往的研究中能够观察到。综上所述,所识别的成像标记物对跟踪AD的进展具有高度的提示性和有效性,与现有的研究结果非常一致。

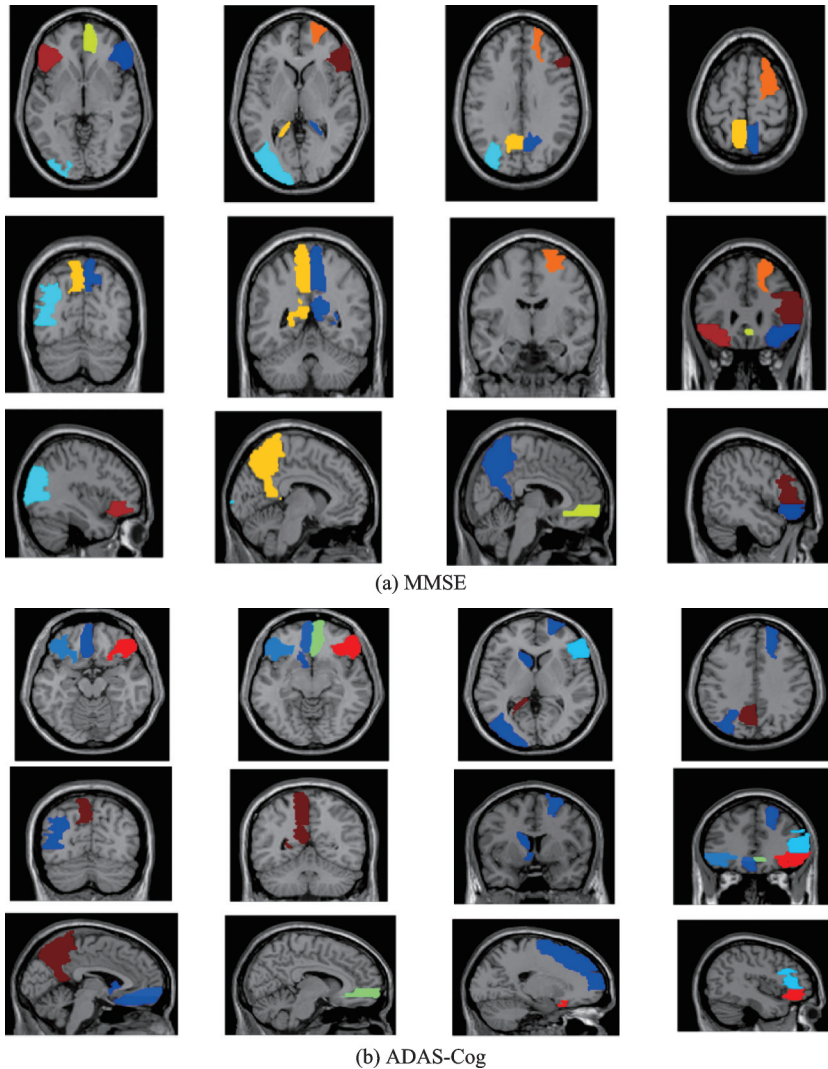


图6 估计临床评分时由本文方法识别的重要脑区

Fig.6 Important brain regions identified by the proposed method when estimating clinical scores

表2 由本文方法识别出的最重要的10个感兴趣区域

Table 2 The most important ten regions of interest identified by the proposed method

MMSE		ADAS-Cog	
ROI index	ROI name	ROI index	ROI name
40	ParaHippocampal_R	40	ParaHippocampal_R
37	Hippocampus_L	37	Hippocampus_L
48	Lingual_R	65	Angular_L
85	Temporal_Mid_L	85	Temporal_Mid_L
42	Amygdala_R	38	Hippocampus_R
65	Angular_L	42	Amygdala_R
86	Temporal_Mid_R	41	Amygdala_L
38	Hippocampus_R	89	Temporal_Inf_L
2	Precentral_R	48	Lingual_R
1	Precentral_L	7	Frontal_Mid_L

4 结束语

本文提出了一种基于参数分解和关系诱导的多任务学习方法,利用多个时间点的sMRI数据以及对应时间的认知评分进行预测。具体来说,通过多任务学习模型能够更好地集成多个时间节点的影像数据,发现不同时间点间的潜在联系。利用参数分解对不同权重施加不同的正则化项,能够识别去除噪声后的共享特征,提高不同时间点共享特征的鲁棒性,进而提高模型的预测性能。关系诱导能够更好地引导知识信息的学习,帮助我们认识到不同时间点之间的真实关系。在数据集上验证了本文方法的有效性,提出的模型与多种主流模型对比均表现出极佳的性能,获得了更出色的结果。本文还对实验结果进行了可视化,并选择影响AD进展中前10个最重要的脑区。如何解决纵向数据中存在的缺失问题是下一步工作需要继续研究的重点。

参考文献:

- [1] FAN Y, RESNICK S M, WU X, et al. Structural and functional biomarkers of prodromal Alzheimer's disease: A high-dimensional pattern classification study[J]. *Neuroimage*, 2008, 41(2): 277-285.
- [2] Alzheimer's Association. 2016 Alzheimer's disease facts and figures[J]. *Alzheimer's & Dementia*, 2016, 12(4): 459-509.
- [3] PETRELLA J R, COLEMAN R E, DORAISWAMY P M. Neuroimaging and early diagnosis of Alzheimer disease: A look to the future[J]. *Radiology*, 2003, 226(2): 315-336.
- [4] ROSEN W G, MOHS R C, DAVIS K L. A new rating scale for Alzheimer's disease[J]. *The American Journal of Psychiatry*, 1984, 141(11): 1356-1364.
- [5] DUCHESNE S, CAROLI A, GEROLDI C, et al. Relating one-year cognitive change in mild cognitive impairment to baseline MRI features[J]. *Neuroimage*, 2009, 47(4): 1363-1370.
- [6] WANG Y, FAN Y, BHATT P, et al. High-dimensional pattern regression using machine learning: From medical images to continuous clinical variables[J]. *Neuroimage*, 2010, 50(4): 1519-1535.
- [7] LIU Mingxia, ZHANG Daoqiang. Pairwise constraint-guided sparse learning for feature selection[J]. *IEEE Transactions on Cybernetics*, 2017, 46(2): 298-310.
- [8] ZHOU Jiayu, LIU Jun, NARAYAN V A, et al. Modeling disease progression via multi-task learning[J]. *Neuroimage*, 2013, 78: 233-248.
- [9] CARUANA R. Multitask learning[J]. *Machine Learning*, 1997, 28: 41-75.
- [10] NIE Liqiang, ZHANG Luming, MENG Lei, et al. Modeling disease progression via multisource multitask learners: A case study with Alzheimer's disease[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2016, 28(7): 1508-1519.
- [11] JIE Biao, LIU Mingxia, LIU Jun, et al. Temporally constrained group sparse learning for longitudinal data analysis in Alzheimer's disease[J]. *IEEE Transactions on Biomedical Engineering*, 2016, 64(1): 238-249.
- [12] CAO Peng, SHAN Xuanfeng, ZHAO Dazhe, et al. Sparse shared structure based multi-task learning for MRI based cognitive performance prediction of Alzheimer's disease[J]. *Pattern Recognition*, 2017, 72: 219-235.
- [13] JIANG Pengbo, WANG Xuetong, et al. Correlation-aware sparse and low-rank constrained multi-task learning for longitudinal analysis of Alzheimer's disease[J]. *IEEE Journal of Biomedical & Health Informatics*, 2018, 23(4): 1450-1456.
- [14] ZHANG Y, YEUNG D Y. A convex formulation for learning task relationships in multi-task learning[C]//*Proceedings of the 26th Conference on Uncertainty in Artificial Intelligence*. [S.l.]: ACM, 2010: 733-742.
- [15] HASTIE T, TIBSHIRANI R J, FRIEDMAN J H. The elements of statistical learning: Springer[J]. *Elements*, 2009, 1(3): 267-268.
- [16] LIU Jun, JI Shuiwang, YE Jieping. Multi-task feature learning via efficient $L_{2,1}$ -norm minimization[C]//*Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*. [S.l.]: ACM, 2009: 339-348.
- [17] WANG Mingliang, ZHANG Daoqiang, SHEN Dinggang et al. Multi-task exclusive relationship learning for Alzheimer's disease progression prediction with longitudinal data[J]. *Medical Image Analysis*, 2019, 53: 111-122.
- [18] DU Lei, LIU Fang, LIU Kefei, et al. Associating multi-modal brain imaging phenotypes and genetic risk factors via a dirty

- multi-task learning method[J]. *IEEE Transactions on Medical Imaging*, 2020, 39(11): 3416-3428.
- [19] WANG Mingliang, HUANG Jiashuang, LIU Mingxia, et al. Functional connectivity network analysis with discriminative hub detection for brain disease identification[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. [S.l.]: AAAI, 2019, 33: 1198-1205.
- [20] ARGYRIOU A, EVGENIOU T, PONTIL M. Convex multi-task feature learning[J]. *Machine Learning*, 2008, 73(3): 243-272.
- [21] STONNINGTON C M, CHU C, KLÖPPEL S, et al. Predicting clinical scores from magnetic resonance scans in Alzheimer's disease[J]. *Neuroimage*, 2010, 51(4): 1405-1413.
- [22] HOSSEINIAN S, AREFIAN E, RAKHSH-KHORSHID H, et al. A meta-analysis of gene expression data highlights synaptic dysfunction in the hippocampus of brains with Alzheimer's disease[J]. *Scientific Reports*, 2020, 10(1): 1-9.
- [23] FENG Qi, NIU Jialing, WANG Luoyu, et al. Comprehensive classification models based on amygdala radiomic features for Alzheimer's disease and mild cognitive impairment[J]. *Brain Imaging and Behavior*, 2021, 15(5): 2377-2386.

作者简介:



郝小可(1985-),男,副教授,硕士生导师,研究方向:机器学习、模式识别、医学图像分析,E-mail:ha-oxiaoke@hebut.edu.cn。



谭麒豪(1999-),男,硕士研究生,研究方向:医学图像分析、模式识别。



李家旺(1998-),男,硕士研究生,研究方向:医学图像分析、模式识别。



郭迎春(1970-),通信作者,女,副教授,硕士生导师,研究方向:图像处理、模式识别,E-mail:gyc@scse.hebut.edu.cn。



于明(1964-),男,教授,博士生导师,研究方向:计算机视觉、人工智能。

(编辑:王静)