

多模态集成阿尔茨海默病和轻度认知障碍分类

程波^{1,2} 钟静¹ 熊江¹

¹(重庆三峡学院计算机科学与工程学院 重庆 404000)

²(南京航空航天大学计算机科学与技术学院 南京 210016)

摘要 为了更有效而准确地诊断阿尔茨海默病 (Alzheimer's disease, AD) 和轻度认知障碍 (Mild Cognitive Impairment, MCI), 文章提出了一种基于多模态数据 (MRI、PET 和非成像数据 CSF) 的集成支持向量机来分类 AD 和 MCI。该算法使用集成学习技术来综合利用不同模态数据之间相互作用产生的分类判别信息, 并利用支持向量机进行分类。为了评价该算法的有效性, 采用十折 (10-fold) 交叉验证策略来验证其性能, 并在标准数据集 ADNI 上测试算法性能。实验结果表明, 多模态集成支持向量机分类方法的性能优于多模态多核学习和单模态方法。

关键词 阿尔茨海默病; 轻度认知障碍; 多模态集成学习; 支持向量机

Multimodal Ensemble Classification of Alzheimer's Disease and Mild Cognitive Impairment

CHENG Bo^{1,2} ZHONG Jing¹ XIONG Jiang¹

¹(Department of Computer Science and Engineering, Chongqing Three Gorges University, Chongqing 404000, China)

²(Department of Computer Science and Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract To effectively diagnose Alzheimer's disease (AD) and mild cognitive impairment (MCI), a multimodal ensemble support vector machine (SVM) based on multi-modality data was proposed and used for the classification of AD and MCI. The ensemble learning was employed and the discrimination information of classification was extracted from different multiple modalities data, then the SVM was used for classification of AD and MCI. In order to validate the efficacy of proposed method, a 10-fold cross-validation was used and tested on the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Experimental results show that the proposed method is better than multi-modality linear multiple kernel learning and single-modality method.

Keywords Alzheimer's disease; mild cognitive impairment; multimodal ensemble learning; support vector machine

1 引言

阿尔茨海默病 (AD) 是发生在老年期及老年前期的一种原发性退行性脑病, 是一种持续性高级神经功能活动障碍。据报道, 全世界大约有 2 千多万阿尔茨海默病患者。对病症早期阶段的准确诊断有助于疾病的治疗。阿尔茨海默病的早期阶段表现为轻度认知障碍 (MCI), 因此, 对 MCI 的准确诊断对阿尔兹海默

病的治疗具有十分重要的意义。根据阿尔茨海默病的病理得知, AD 的产生与发展是一个缓慢并持续的过程。MCI 经过多年可能会发展成为 AD, 也可能不发生变化。在阿尔茨海默病早期诊断阶段, 不少研究者采用神经影像的技术 (如: 核磁共振成像 (MRI)) 来诊断病情, 并获得了较好的效果。当前的模式分类方法对 AD 与 HC (Healthy Controls) 分类效果较好, 但是分类 MCI 与 HC 效果不太理想^[1-4]。最近一些研究表明, 采用多模态的生物标记对诊断 AD 和 MCI 效果较

基金项目: 重庆市教委科学技术研究项目 (No. KJ121111 和 KJ131108)。

作者简介: 程波 (通讯作者), 博士研究生, 讲师, 研究方向为机器学习、医学图像分析, E-mail: chengbo729@126.com; 钟静, 副教授, 研究方向为计算智能、神经网络在计算机网络中的应用研究; 熊江, 教授, 研究方向为软件工程、数据挖掘与医学图像分析。

好,其中包括测量脑萎缩的核磁共振成像(MRI),正电子断层成像(PET)和非成像数据(CSF)^[5-7]。其中有采用多核学习技术融合多模态数据的AD或MCI分类方法,且都优于单模态数据的方法^[1,2,4]。

当前研究表明,基于单模态数据的AD或MCI分类方法较多,少数研究工作是基于多模态数据。在基于多模态数据的研究工作中,大多数采用线性多核学习方法来融合多模态数据,再使用支持向量机(SVM)等分类方法分类^[1,2,4]。基于线性多核学习的多模态分类方法能得到比单模态分类方法更好的分类性能。但是,利用线性多核技术提取的决策信息只来源于多个模态数据,而不能充分地利用各个模态数据之间相互作用对分类问题产生互补的决策信息。参考文献[8]中提出了一种线性支持向量机集成分类方法,用于诊断AD和自闭症障碍(Autism Spectrum Disorder, ASD),其分类性能较好。因此,本文提出了一种基于多模态数据的集成支持向量机方法用于诊断AD和MCI。具体步骤为:首先分别对MRI和PET图像进行预处理,并提取特征;再使用t-test方法分别对提取的MRI和PET特征进行选择;接着分别使用线性SVM对MRI、CSF和PET训练分类器;最后对这些分类器进行集成。为了评价该算法的有效性,采用10-fold交叉验证策略来计算分类精度(Accuracy)、敏感度(Sensitivity)、特异度(Specificity)和AUC值。在标准数据集ADNI上的实验结果表明,本文提出的多模态数据的集成支持向量机方法的分类性能优于多模态线性多核学习方法和单模态方法。

2 基于多模态数据的集成支持向量机

基于多模态数据的集成支持向量机方法在本质上是支持向量机决策函数在不同模态数据上的线性组合。由于不同模态数据对脑疾病诊断能提供互补的诊断信息,因此,本文提出了集成支持向量机方法,综合使用不同模态数据互补的脑疾病诊断信息。在本节中,首先介绍标准支持向量机算法,然后将标准支持向量机算法扩展到集成支持向量机。

2.1 支持向量机

支持向量机分类的核心思想是:针对两类问题,能够找到一个使样本点尽量分开,同时使分类样本到这个超平面的距离尽可能大的最优超平面。其可通过一个带约束的凸规划问题求解。设已知两类问题

的训练集T为 $\{x_i, y_i\}_{i=1}^n$,其中 $x_i \in R^D$ 是D维样本, $y_i \in \{+1, -1\}$ 是样本 x_i 对应的类别标签。寻找在已知训练集T上的一个决策函数 $f(x) = \langle w, \Phi(x) \rangle + b$,其中 $\Phi: R^D \rightarrow R^d$ 的映射函数,并能推断任一输入样本x相对应的类别y,也就是求解一个把 R^D 上的点分成两部分的规则。其目标就是如下优化问题:

$$\begin{aligned} \min_{w, C, \xi_i} & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \\ \text{s. t. } & y_i (\langle w, \Phi(x_i) \rangle + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i=1, \dots, n \end{aligned} \quad (1)$$

其中w是权值系数,C是惩罚参数。(1)式是支持向量机分类求解构造凸二次规划的原始问题,利用拉格朗日乘子法把原始问题转化为其对偶问题的求解,则(1)式对偶问题为:

$$\begin{aligned} \max_{\alpha} & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \langle \Phi(x_i), \Phi(x_j) \rangle \\ \text{s. t. } & \mathbf{k}(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle; \sum_{i=1}^n \alpha_i y_i = 0 \\ & C \geq \alpha_i \geq 0, i=1, \dots, n \end{aligned} \quad (2)$$

其中 $\mathbf{K} = \mathbf{k}(x_i, x_j)$ 是核函数, α 是拉格朗日乘子。

而(2)式中的优化问题为凸规划问题,进行求解得最终决策函数为:

$$f(x) = \sum_{i=1}^n \alpha_i y_i \mathbf{k}(x_i, x) + b \quad (3)$$

2.2 多模态集成支持向量机

利用集成学习技术,将标准支持向量机扩展到集成支持向量机。首先,对于多模态数据,其集成的决策函数定义为:

$$F(x) = \sum_{m=1}^M c_m f^{(m)}(x^{(m)}) \quad (4)$$

其中 c_m 是权值,并满足 $\sum_{m=1}^M c_m = 1$, $f^{(m)}$ 表示m-th模态数据的支持向量机决策函数, $F(x)$ 为最终决策函数。

不同模态的数据对疾病诊断具有不同的效果,因此,需采用不同的权值来体现这种差别。由于使用不同单模态数据方法将得到不同的分类精度,采用单模态方法分类精度所占的比重作为该模态在多模态数据的集成支持向量机方法中的权值,能体现不同模态数据对分类结果的贡献度。所以,对于权值 c_m 的计算,采用(5)式的方法获得。则权值 c_m 为:

$$c_m = \frac{Accuracy_m^{f^{(m)}}}{\sum_{m=1}^M Accuracy_m^{f^{(m)}}} \quad (5)$$

其中, $Accuracy_m^{f^{(m)}}$ 表示基于m-th模态数据支持向量机的分类精度。结合(4)和(5)式,就能求解最终决策函数。

3 实验与结果分析

本文提出的方法基于 MRI、PET 和 CSF 三种模态数据, 对 AD 与 HC、MCI 与 HC 两组情况进行分类。实验所用数据来自于标准数据集 ADNI (www.loni.ucla.edu/ADNI)。所使用数据的被试者信息(均值±标准差)如表 1。由于标准数据集 ADNI 中同时具有三种模态数据的被试者只有 202 个。所以, 本文采用这 202 个被试者的数据用于实验。

表 1 被试者信息

| | AD ($n=51$) | MCI ($n=99$) | HC ($n=52$) |
|-------|---------------|----------------|---------------|
| 年龄 | 75.2±7.4 | 75.3±7.0 | 75.3±5.2 |
| 受教育程度 | 14.7±3.6 | 15.9±2.9 | 15.8±3.2 |

3.1 图像分析与实验设置

对 ADNI 数据集上的 MRI 和 PET 图像采用参考文献[1]中的图像预处理方法进行预处理。具体过程为: 首先采用 N3 算法^[9]对所有图像进行前连合与后连合校正, 再使用 BSE (Brain Extraction Tool)^[10]工具包将颅骨剥离, 去除小脑。然后使用 FSL 工具包^[11]把 MRI 图像分割成三个不同的组织: 灰质、白质和脑脊髓液。接着使用 HAMMER 算法^[12]对所有图像进行配准, 并标注出所有 MRI 和 PET 图像的 93 个感兴趣区域。再计算出每个被试者的 93 个感兴趣区域的 MRI 图像灰质组织体积 (Tissue Volume)。而对 PET 图像计算出每个被试者的每个感兴趣区域组织体积的均值。经过预处理后得到的特征是: 来自 MRI 图像的 93 个特征, 来自 PET 图像的 93 个特征和来自非图像数据 CSF 的 3 个特征 (分别是: CSF $A\beta_{42}$, CSF t -tau 和 CSF p -tau)。最后, 分别用 t -test 方法对 MRI 和 PET 数据进行特征选择。

为了有效评价算法的分类性能, 本文采用十折交叉验证的方法计算分类器的精度、敏感度和特异度。具体过程为: 数据集被均匀的随机划分成 10 份, 依次选择其中 1 份作为测试集, 其余 9 份为训练集, 重复 10 次十折交叉验证实验, 最终使用全部精度、敏感度和特异度的均值作为分类器的精度、敏感度和特异度。支持向量机分类算法采用的是 LIBSVM 工具包, 每个模态数据都采用线性核函数, 惩罚参数 ($C+1$) 等其他参数都使用默认值。 t -test 特征选择方法 p 的选择和权值 c_m 的确定分别在训练集上交叉验证获得。MRI、CSF 和 PET 模态数据标准化与参考文献[1]中的方法相同。

3.2 实验结果与分析

为了验证多模态集成支持向量机分类 AD 和 MCI 的有效性, 分别采用单模态与多模态方法来进行比较。表 2 和表 3 分别是分类 AD 和 MCI 在单模态与多模态方法上比较实验的结果。其中, MRI、CSF 和 PET 表示单模态的 SVM 分类方法, 而 MRI+PET+CSF 表示采用三个模态数据的多模态集成 SVM 分类方法, AUC 表示 ROC 曲线与横坐标轴之间区域的面积。图 1 和图 2 分别是分类 AD 和 MCI 在单模态与多模态方法上的 ROC 曲线。表 2、表 3 和图 1、图 2 的实验结果表明, 对于分类 AD 和 MCI, 多模态集成 SVM 分类方法都明显优于单模态方法。对于分类 AD 与 HC, 多模态方法能获得精度为 96.2%、敏感度 98%、特异度 94.6% 和 AUC 为 0.986 的高识别率; 而单模态方法最好情况是使用 PET 数据, 能获得精度为 86.5%、敏感度 86.3%、特异度 86.6% 和 AUC 为 0.938 的识别率。对于分类 MCI 与 HC, 多模态方法能获得精度为 80.9%、敏感度 87.1%、特异度 70.3% 和 AUC 为 0.834 的识别率; 而单模态方法最好情况是使用 MRI 数据, 能获得精度为 72%、敏感度 78.5%、特异度 59.6% 和 AUC 为 0.759 的识别率。以上实验结果表明: 多模态方法能充分利用各个模态数据对疾病诊断的互补信息, 能完成比单模态方法更好的分类性能。

表 2 AD 与 HC 分类在单模态和多模态方法上的性能

| 方法 | 精度 | 敏感度 | 特异度 | AUC |
|-------------|-------|-------|-------|-------|
| MRI | 86.2% | 86.0% | 86.3% | 0.936 |
| CSF | 82.1% | 81.9% | 82.3% | 0.887 |
| PET | 86.5% | 86.3% | 86.6% | 0.938 |
| MRI+PET+CSF | 96.2% | 98.0% | 94.6% | 0.986 |

表 3 MCI 与 HC 分类在单模态和多模态方法上的性能

| 方法 | 精度 | 敏感度 | 特异度 | AUC |
|-------------|-------|-------|-------|-------|
| MRI | 72.0% | 78.5% | 59.6% | 0.759 |
| CSF | 71.4% | 78.0% | 58.8% | 0.762 |
| PET | 71.6% | 78.2% | 59.3% | 0.744 |
| MRI+PET+CSF | 80.9% | 87.1% | 70.3% | 0.834 |

为了验证多模态集成支持向量机融合多模态数据的有效性, 分别采用线性融合多模态数据方法 (参考文献[1]中 Zhang 等提出的方法) 和投票方式集成 SVM 方法 (表 4 中的 “Vote”) 来比较本文提出的方法。表 4 是三种多模态方法的分类性能比较, 表中 “本文” 表示本文提出的多模态非线性多核支持向量机, “Zhang” 表示参考文献[1]中提出的多模态线性多核支持向量机, “Vote” 表示投票方式集成 SVM 方法。图 1 和图 2 分别绘制了三种方法的 ROC 曲线。表 4 和图 1、图 2 和表 4 表明, 本文提出的方法

优于参考文献[1]中 Zhang 等提出的方法和“Vote”方法。图 1 和图 2 也表明, 三种多模态方法的分类性能都明显优于单模态的分类方法; 而本文提出的方法均优于以上所有方法。这说明, 采用多模态集成支持向量机能有效融合多模态数据并改善了分类 AD 和 MCI 的性能。

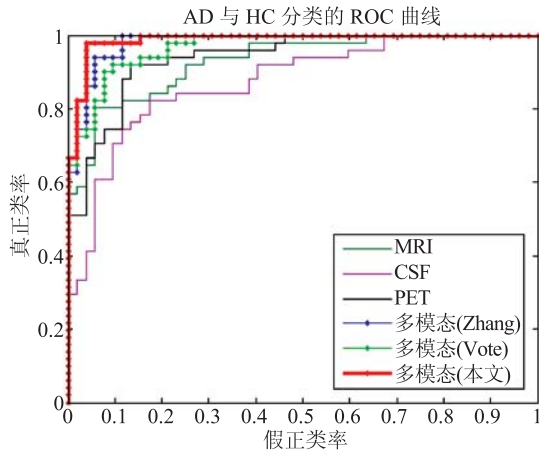


图 1 六种方法对 AD 与 HC 分类的 ROC 曲线

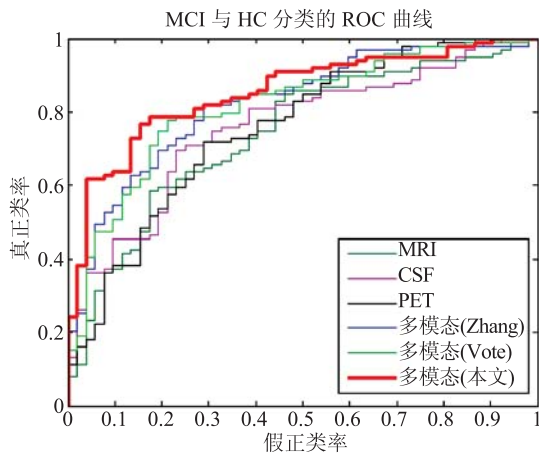


图 2 六种方法对 MCI 与 HC 分类的 ROC 曲线

表 4 三种多模态方法比较

| 方法 | 精度 | 敏感度 | 特异度 | AUC |
|------------|-------|-------|-------|-------|
| AD vs. HC | | | | |
| 本文 | 96.2% | 98.0% | 94.6% | 0.986 |
| 51AD+52HC | | | | |
| Zhang | 93.2% | 93.0% | 93.3% | 0.976 |
| Vote | 91.2% | 90.6% | 91.8% | 0.966 |
| MCI vs. HC | | | | |
| 本文 | 80.9% | 87.1% | 70.3% | 0.834 |
| 99MCI+52HC | | | | |
| Zhang | 76.4% | 81.8% | 66.0% | 0.809 |
| Vote | 77.6% | 85.9% | 62.1% | 0.822 |

4 结束语

本文根据集成学习能提取不同模态数据相互作用产生互补的分类判别信息的原理, 采用多模态集成支持向量机和多模态数据来分类 AD 和 MCI。在标准数

据集 ADNI 上的实验结果表明: 本文提出的多模态集成支持向量机的分类性能优于多模态线性多核支持向量机、投票方式集成支持向量机和单模态方法。该方法能有效利用不同模态数据相互作用产生互补的分类判别信息, 体现了多模态集成支持向量机融合多模态数据方法的有效性。下一步的工作将把该方法从两类分类推广到多类分类。

参考文献

- [1] Zhang D, Wang Y, Zhou L, et al. Multimodal classification of Alzheimer's disease and mild cognitive impairment [J]. NeuroImage, 2011, 55(3): 856-867.
- [2] Hinrichs C, Singh V, Xu G, et al. Predictive markers for AD in a multi-modality framework: an analysis of MCI progression in the ADNI population [J]. NeuroImage, 2011, 55(2): 574-589.
- [3] Davatzikos C, Fan Y, Wu X, et al. Detection of prodromal Alzheimer's disease via pattern classification of magnetic resonance imaging [J]. Neurobiology of Aging, 2008, 29(4): 514-523.
- [4] Ye J, Chen K, Wu T, et al. Heterogeneous data fusion for Alzheimer's disease study [C] // Proceedings of the 14th ACM International Conference on Knowledge Discovery and Data Mining, 2008: 1025-1033.
- [5] Fjell AM, Walhovd KB, Fennema NC, et al. CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease [J]. Journal of Neuroscience, 2010, 30(6): 2088-2101.
- [6] Foster NL, Heidebrink JL, Clark CM, et al. FDG-PET improves accuracy in distinguishing front temporal dementia and Alzheimer's disease [J]. Brain, 2007, 130(10): 2616-2635.
- [7] Du AT, Schuff N, Kramer JH, et al. Different regional patterns of cortical thinning in Alzheimer's disease and front temporal dementia [J]. Brain, 2007, 130(4): 1159-1166.
- [8] Varol E, Gaonkar B, Erus G, et al. Feature ranking based nested support vector machine ensemble for medical image classification [C] // Proceedings of the International Symposium on Biomedical Imaging, 2012: 146-149.
- [9] Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity non-uniformity in MRI data [J]. IEEE Transactions on Medical Imaging, 1998, 17(1): 87-97.
- [10] Smith SM. Fast robust automated brain extraction [J]. Human Brain Mapping, 2002, 17(3): 143-155.
- [11] Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation maximization algorithm [J]. IEEE Transactions on Medical Imaging, 2001, 20(1): 45-57.
- [12] Shen D, Davatzikos C. HAMMER: hierarchical attribute matching mechanism for elastic registration [J]. IEEE Transactions on Medical Imaging, 2002, 21(11): 1421-1439.