

正则化多任务学习在精神分裂症核磁共振成像图像分类中的应用

张娜,王瑜,周文,肖洪兵,邢素霞

北京工商大学计算机与信息工程学院食品安全大数据技术北京市重点实验室,北京 100048

【摘要】为实现对精神疾病的计算机辅助诊断与预后,利用机器学习与图像处理技术分析多地区精神疾病的核磁共振成像数据,已成为该领域的必然趋势。本文首先提出切片提取的核磁共振成像图像预处理方法,然后提取图像的纹理特征,最后提出一种 l_p 范数正则化的多任务学习支持向量机精神分裂症分类方法,同时学习3个数据中心精神分裂症图像的共享特征和各自独有的特征,用于分类精神分裂症患者和正常人。实验结果表明,该方法取得了优秀的诊断精度,可为精神分裂症患者的临床诊断与治疗提供生物学依据。

【关键词】精神分裂症;磁共振成像;特征提取;正则化多任务学习

【中图分类号】TP391.9;R445.2

【文献标志码】A

【文章编号】1005-202X(2018)07-0790-06

Application of regularized multi-task learning in schizophrenia MRI data classification

ZHANG Na, WANG Yu, ZHOU Wen, XIAO Hongbing, XING Suxia

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China

Abstract: Machine learning techniques and magnetic resonance imaging (MRI) techniques have been used in the analysis of MRI data of patients with mental diseases in various regions to achieve computer-aided diagnosis and prognosis of mental diseases such as schizophrenia, etc. Herein slice extraction is firstly used for MRI image preprocessing. Then texture features of gray-level co-occurrence matrices are extracted from the above processed images. Finally, a l_p -norm regularized multi-task learning method based on support vector machine for MRI data classification is proposed to simultaneously learn the site-specific and site-shared features of schizophrenia images from 3 data centers, which can be used to discriminate schizophrenia patients from normal controls. Experiments show that the proposed method achieves a high diagnosis accuracy, providing a biological basis for the clinical diagnosis and treatment of schizophrenia.

Keywords: schizophrenia; magnetic resonance imaging; feature extraction; regularized multi-task learning

前言

伴随着科技的发展,各类高端医学影像设备的发展日趋迅速,医学影像在协助医生诊断疾病方面发挥着越来越重要的作用。但大量的医学影像明显增加了医生的工作负担,因此,如何借助机器学习和图像处理技术,让计算机辅助医生诊断治疗,是人工

智能医疗的研究热点之一。在众多的医学影像中,核磁共振成像(MRI)技术以其无辐射、分辨率高等优点^[1],已被广泛应用到脑疾病的临床诊断中。目前,研究者多根据单一地区某种脑疾病的MRI图像,进行图像分割、识别、分类等研究工作,因此研究者往往忽略了对多地区同种脑疾病的MRI图像的相关研究。然而,在互联网信息化时代,通过多种渠道获取多地区同种脑疾病的MRI图像已成为可能,文献[2-4]研究表明,与单一地区的少量样本相比,多地区的同种脑疾病的MRI图像数据能够提供更为充足的统计样本,从而能更好探索病人的大脑结构功能模式。文献[5-8]研究发现,与单一地区的疾病人群相比,同一种疾病在不同地区的人口分布差异是多样化的,如不同人种的大脑结构功能存在一定差异,多地区的人群样本患病严重程度及出现的临床症状不

【收稿日期】2018-02-10

【基金项目】国家自然科学基金(61671028);北京市自然科学基金(4162018);北京市委组织部“高创计划”青年拔尖人才培养资助项目(2014000026833ZK14);北京市青年拔尖人才培养计划项目(CIT&TCD201504010)

【作者简介】张娜,硕士研究生,研究方向:医学图像处理、机器学习, E-mail: zhangna1073@163.com

【通信作者】王瑜,博士,副教授,硕士生导师,研究方向:医学图像处理、模式识别, E-mail: wangyu@btbu.edu.cn

同,患者类型更广泛等。显然,研究多地区的同种疾病人群的医学图像,不仅能获得更为全面的图像信息或一致的异常病理特征模式,还能分析单一地区内该疾病医学图像特点,实验结果更具说服力。同时,随着国际医疗机构的广泛合作及医工交叉,采用多地区患病人群的医学影像数据来研究该病的病理机制,是一种必然趋势。

由上可知,若研究精神分裂症脑疾病的多地区MRI图像数据,借助计算机辅助诊断技术来区分正常、异常MRI图像,最终将患者和正常人正确分类,其优势显而易见。但任何分类器都是在有限的样本条件下训练出来的,尤其是在MRI图像样本规模小、多样性少的情况下,很难用有限的样本模式代替无限的样本模式,也达不到与实际模式的高度吻合。针对这一问题,多任务学习方法的优势逐渐体现出来,多任务学习是一种通过挖掘任务间的共享信息,同时训练多个相关任务的模型,从而获得最优的学习方法。它可以显著地提高算法的学习效果,并已经应用到多个领域,如垃圾邮件的过滤^[9]、自然图像的分类^[10]、各种疾病建模分类和预测^[11-12]等。文献[13-14]表明,多任务之间通过相互归纳偏置的作用,获得偏置信息补充样本的不足,能够同时学习单任务独有的特征信息和多任务间共享的特征信息,有效提高模型泛化能力。Evgeniou等^[15]提出正则化多任务学习方法,该方法基于支持向量机(Support Vector Machine, SVM)模型,通过添加正则化惩罚项,约束不同任务模型参数的相关性,从而提高模型的泛化能力。刘建伟等^[16]提出一种 l_1 、 l_2 混合范数正则化多任务学习SVM方法,并应用于抑郁症疾病患者的MRI图像分类,取得较高的分类准确率。

在前人研究基础上,本研究将多数据中心的精神分裂症MRI图像分类问题看成是一个多任务学习问题,构建了 l_p 范数正则化多任务学习SVM分类模型,并采用梯度下降法优化模型,最后将模型应用在正常和异常MRI图像的分类上,从而实现了对患者的正确分类。

1 研究方法

1.1 图像预处理

MRI数据通常以三维数据样本的形式存储,本实验中,所用结构磁共振成像(structural Magnetic Resonance Imaging, sMRI)灰质图像大小为 $96 \times 113 \times 94$,若直接对每个体素进行特征提取,会出现维度灾难的问题,大量无关或是冗余的特征信息极大降低了模型性能。因此,本文提出一种切片化并计算加

权求和平均灰度图像的预处理方法,其步骤如下:(1)切片化,每个人的灰质图像大小为 $96 \times 113 \times 94$,沿Z轴方向进行切片化处理,得到94片切片。(2)筛选并转换为灰度图像,去掉灰质图像中无特征信息量的10个切片,将剩余的各个切片转换成灰度图像,并依次编号为 $a_i (i=0, 1, 2, \dots, 83)$ 。sMRI图像库中编号为NC001的部分灰质图像切片灰度化后如图1所示。(3)加权求平均,根据各个切片包含的特征信息量,赋予不同的权重,按顺序均分成3组,分别计算各组中加权求和平均灰度图像Img1、Img2、Img3,各组计算公式分别为:

$$\text{Img1} = \frac{m_0 * 1 + m_1 * 2 + m_2 * 3 + \dots + m_{27} * 28}{1 + 2 + 3 + \dots + 28} \quad (1)$$

$$\text{Img2} = \frac{m_{28} * 1 + m_{29} * 1 + m_{30} * 1 + \dots + m_{55} * 1}{28} \quad (2)$$

$$\text{Img3} = \frac{m_{29} * 28 + m_{30} * 27 + m_{31} * 26 + \dots + m_{83} * 1}{28 + 27 + 26 + \dots + 2 + 1} \quad (3)$$

完成预处理后,每个样本包含Img1、Img2、Img3这3幅图像,为后续特征提取做好准备工作。

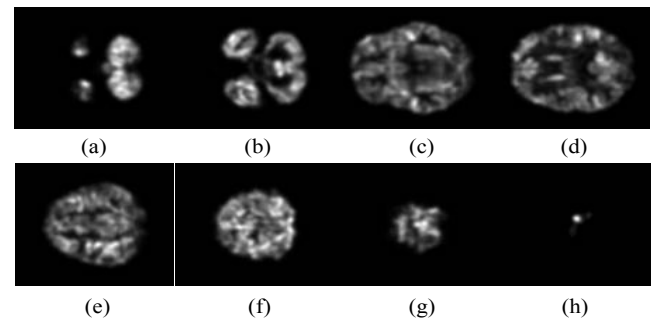


图1 部分灰质图像实例

Fig.1 Partial gray matter image example

(a) to (h) were the 13th, 20th, 34th, 48th, 62th, 69th, 76th and 83th slices, respectively

1.2 特征提取

图像的特征提取是计算机处理医学图像的一个基本且关键的步骤,目的是将样本固有的、本质的、区别于其他样本的特征或属性以特征矢量等形式表示出来,其结果直接影响分类准确度。由于MRI图像中的纹理信息对噪声、光照和颜色不敏感,因此,本文提取基于灰度共生矩阵的纹理特征进行研究。

灰度共生矩阵是一个统计图像中相邻(或一定间距内)两像素的灰度呈某种关系的矩阵,反映图像灰度在方向、间隔、变化幅度上的综合信息^[14]。设一幅二维数字图像灰度级为 N , $p(i, j)$ 表示给定空间距离为 d ,方向角为 θ ,灰度级以 i 为始点,出现灰度级为 j 的概率(或频数)。灰度共生矩阵提供图像的灰度方向、间

隔和变化幅度的统计信息,依据这些统计信息,可以计算以下用来定量描述纹理特性的统计量:(1)均值

$$\text{Mean} = \bar{x} = \sum_{i=0}^N \sum_{j=0}^N p(i,j) \times i, \text{均值反映纹理的规则程度,}$$

纹理越杂乱无章,或难以描述,其值越小,反之越大;(2)

$$\text{Variance} = \sum_{i=0}^N \sum_{j=0}^N p(i,j) \times (i - \bar{x})^2, \text{方差反映象元值}$$

与均值偏差的度量,当图像中灰度变化较大时,方差值

$$\text{Entropy} = - \sum_{i=0}^N \sum_{j=0}^N p(i,j) \times \ln p(i,j), \text{熵用来度}$$

量图像具有的信息量,是测量灰度级分布随机性的特征参数,表征了图像中纹理的复杂程度,纹理越复杂熵

$$\text{Contrast} = \sum_{i=0}^N \sum_{j=0}^N p(i,j) \times (i-j)^2, \text{对比度}$$

反映图像中局部灰度变化总量,局部像素对的灰度差别越大,图像的对比度越大,图像视觉效果越清晰;

(5) 相关性

$$\text{Correlation} = \sum_{i=0}^N \sum_{j=0}^N \frac{(i - \text{Mean}) \times (j - \text{Mean}) \times p(i,j)}{\text{Variance}}, \text{相关性是}$$

灰度线性关系的度量,反映某种灰度值沿某方向的延伸长度,延伸越长,相关性越大;(6)同质度(或逆差矩)

$$\text{Homogeneity} = \sum_{i=0}^N \sum_{j=0}^N p(i,j) \times \frac{1}{1 + (i-j)^2}, \text{同质度用来度}$$

量图像局部灰度的均匀性,局部灰度越均匀,同质度值

$$\text{Energy} = \sum_{i=0}^N \sum_{j=0}^N p(i,j)^2, \text{能量(或角二阶矩)}$$

用来度量图像灰度分布的均匀性,当灰度共生矩阵

(Grey-level Co-occurrence Matrix, GLCM)中元素分布集中于主对角线附近时,说明局部区域内的图像灰度

分布均匀,能量值相应较大,当GLCM的所有值相等时,能量值较小。

综上所述,对每个被试的三幅加权求和平均灰度图像(Img1、Img2、Img3)分别计算上述7个统计量,每个样本共得到具有21个属性的特征向量。为描述统一样本特征值的统计分布特性,需利用公式 $x'_i = \frac{x_i - \mu}{\sigma_x}$ 进行归一化处理,其中, x_i 为特征值, μ 为所有样本数据的均值, σ 为所有样本数据的标准差。

1.3 分类算法

1.3.1 多任务学习与单任务学习 机器学习算法通常是一次学习一个任务,将复杂的学习问题分解为理论上独立的子问题,然后分别对每个子问题进行学习,最后对子问题的学习结果的组合建立复杂问题的数学模型,即单任务学习的思想^[17-18]。多任务学习是和单任务学习相对的一种机器学习方法,它利用任务间共享的信息同时对多个任务并行学习,同时求解多个问题,且结果互相影响。任务间存在共享信息是进行多任务学习的前提,在此基础上,多个任务的训练才可以提升模型的整体泛化性能。

多任务学习与单任务学习的主要区别在于模型的训练过程不同^[19-20]。单任务学习的训练中,每个任务都是独立进行的,互不影响,其弊端在于单独进行一项任务训练时,忽视了训练其他任务所包含的信息,某种程度上造成相关训练信息的丢失,而且这部分信息对训练过程来说是非常有用的。而多任务学习的训练中考虑了任务间的相关性和共享的有用信息,同时并行地学习多个任务,二者在训练模型过程中的区别如图2所示。

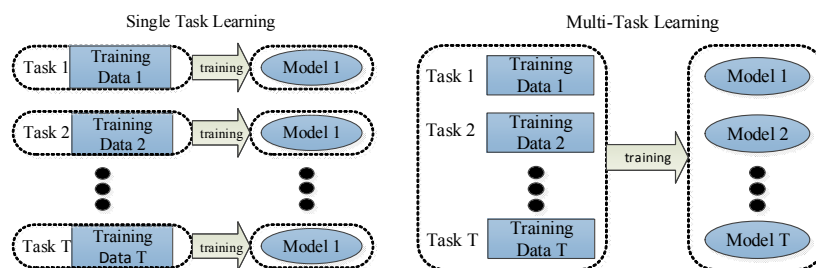


图2 单任务学习和多任务学习模型训练过程对比

Fig.2 Comparison of the model training process of single task learning and multi-task learning

1.3.2 正则化多任务SVM 多任务学习最重要的问题是如何对各任务间的关系进行建模,并使相关任务可以共享信息,最终达到利用不同任务间的相关性提升算法学习性能的目标,期望模型尽量拟合训练数据,同时又要避免模型过于复杂,因此采用添加正则化项的多任务SVM算法,解决多数据中心精神

疾病的MRI图像分类问题。

假设多任务学习问题中共有 t 个监督学习任务,对于其中任意一个任务 i ,假定其学习函数为 $f_i: \mathbf{R}^d \rightarrow \mathbf{R}$, 训练集为 $\mathbf{X}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbf{R}^{d \times n}$, $i = 1, 2, \dots, t$, 其中 n 为输入样本数据的个数, d 为样本特征向量的维数, $\mathbf{Y}_i = [y_1, y_2, \dots, y_n] \in \mathbf{R}^n$, $i = 1, 2, \dots, t$, 其中 $y_i \in \{+1, -1\}$, 为

第 i 个任务中每个样本的类别标签, t 个监督学习任务的权重系数矩阵为 $\mathbf{W}=[w_1, w_2, \dots, w_t] \in \mathbf{R}^{d \times t}$ 。多任务学习的目标是通过训练数据的学习,得到 t 个相关任务的回归或分类函数 $f_i(x)$,为精确求出 $f_i(x)$ 函数,首先应该确定多任务的目标函数。假定 $f(\mathbf{w}_i^T \mathbf{X}_i, \mathbf{Y}_i)$ 为第 t 个任务的损失函数,在分类问题中,经典的损失函数包括对数似然函数、指数函数和Hinge函数,学习一个经验误差最小,并加入正则化的多任务学习SVM模型,可用如下公式^[16]表示:

$$\text{Obj}(w, x, y) = \min_w \frac{1}{n} \sum_{i=1}^n f(\mathbf{w}_i^T \mathbf{X}_i, \mathbf{Y}_i) + \lambda \Omega(\mathbf{W}) \quad (4)$$

其中,上式第一项为训练数据上的经验损失函数, $f(\mathbf{w}_i^T \mathbf{X}_i, \mathbf{Y}_i)$ 采用Hinge损失函数,上式第二项为正则化项,能够编码任务间的相关性, $\lambda > 0$ 为正则化项参数。

单独求解 t 个任务的最优解相当于求解联合 t 个任务的目标函数得到的全局问题:

$$\text{Obj}(w, x, y) = \min_w \sum_{i=1}^t \frac{1}{n} \sum_{i=1}^n f(\mathbf{w}_i^T \mathbf{X}_i, \mathbf{Y}_i) + \lambda \sum_{i=1}^t \Omega(\mathbf{W}) \quad (5)$$

机器学习通常采用模型参数向量的范数作为正则化项,正则化阶次需要提前设定,常用的有 l_0 、 l_1 、 l_2 范数。在本实验中,发现对于不同的数据,使用不同的正则化阶次,可以改进分类算法的分类准确率,故提出 l_p 范数正则化多任务学习SVM分类算法。 l_p 范数理论简单,在处理图像数据的过程中不但有效,而且易于优化,可以降低模型的计算复杂度,其计算公式^[14]如下:

$$h(\mathbf{w}) = \|\mathbf{w}\|_p = \left(\sum_{i=1}^n |w_i|^p \right)^{\frac{1}{p}} \quad (6)$$

其中, $x = \{x_1, \dots, x_n\}$ 为一个向量, l_p 范数度量向量的稀疏性,阶次 p 的可取范围为 $0 < p \leq 2$, P 的选择根据任务间的相关性大小来决定,当任务之间相关性比较大,需要共享的信息较多, P 值应该适当增大,反之选择较小的值。令 $k(\mathbf{w}) = h^p(\mathbf{w})$,当 $0 < P \leq 2$ 时,其导数公式如下:

$$\frac{k(\mathbf{w})}{\partial w_i} = p |w_i|^{p-1} \times \text{sgn}(w_i) \quad (7)$$

由于 $\text{sgn}(w_i) = \frac{w_i}{|w_i|}$,所以式(7)可记为:

$$\frac{k(\mathbf{w})}{\partial w_i} = p |w_i|^{p-2} \times w_i \quad (8)$$

最终,加入 l_p 范数正则项的多任务学习SVM目标函数如式(9)所示:

$$L = \min_w \sum_{i=1}^t \frac{1}{n} \sum_{i=1}^n \max(0, 1 - y_i \mathbf{w}_i^T x_i) + \lambda \sum_{i=1}^t \|\mathbf{w}_i\|_p \quad (9)$$

根据不同的情况,对式(9)求导为:

(a)如果 $1 - y_i \mathbf{w}_{t,i}^T x_i < 0$,则:

$$\frac{\partial L}{\partial w_{t,i}} = 0 \quad (10)$$

(b)如果 $1 - y_i \mathbf{w}_{t,i}^T x_i > 0$,则

$$\frac{\partial L}{\partial w_{t,i}} = -y_i x_{t,i} + p |w_{t,i}|^{p-2} \times w_{t,i} \quad (11)$$

利用梯度下降法对第 t 个任务 $\mathbf{W}_{t,i} = \{\omega_{t,1}, \omega_{t,2}, \dots, \omega_{t,k}\}$ 进行更新:

$$\mathbf{W}_{t,i} = \mathbf{W}_{t-1,i} + r \nabla L(\mathbf{W}_{t-1,i}) \quad (12)$$

2 实验结果与分析

实验环境选用Intel Core i5 2.40 GHz的处理器和32g内存的PC机,开发工具为Matlab2013a、Python2.7,使用Python编程语言。实验中,精神分裂症sMRI图像的特征提取与分类算法流程图如图3所示。

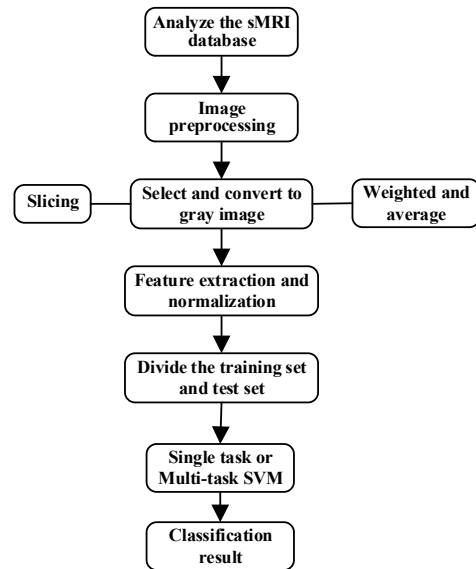


图3 精神分裂症sMRI图像的特征提取与分类算法流程图

Fig.3 Schizophrenia sMRI image feature extraction and classification algorithm flow chart

sMRI: Structural magnetic resonance imaging; SVM: Support vector machine

2.1 图像库介绍

图像库受试者由美国(A地区)的132名正常人和137名精神分裂症患者、巴西(B地区)的94名正常人和62名精神分裂症患者、中国(C地区)的181名正常人和144名精神分裂症患者组成。所有的精神分裂症患者都符合美国精神疾病的诊断与统计手册第四版中精神分裂症的诊断标准(DSM-IV诊断),并且他们都没有其他神经疾病或严重的药物疾病等历史。本研究从每个地区随机选取正常人和患者各60

名进行研究,同时对各地区被试者的相关信息进行统计学分析,结果表明,患者在性别、年龄等方面都

与健康对照组相匹配,差异无统计学意义($P<0.05$),具体如表1所示。

表1 被试者统计学分析结果
Tab.1 Characteristics of the participants in this study

Region	Class	Sample size	Gender (male/female)	Average age/years	Age range/years
A	SCZ	60	37/23	34.68	18-60
	NC	60	28/32	31.78	13-65
	<i>P</i> value	--	0.09 ^a	0.22 ^b	--
B	SCZ	60	44/16	27.53	18-50
	NC	60	37/23	29.73	18-50
	<i>P</i> value	--	0.17 ^a	0.13 ^b	--
C	SCZ	60	23/37	30.85	16-54
	NC	60	28/32	33.93	20-57
	<i>P</i> value	--	0.36 ^a	0.11 ^b	--
A+B+C	SCZ	180	105/75	31.39	16-60
	NC	180	89/91	30.90	15-65
	<i>P</i> value	--	0.09 ^a	0.65 ^b	--

^a: Pearson Chi-square test; ^b: Two-sample *t*-test; SCZ: Schizophrenics; NC: Normal controls

2.2 实验设计

为验证本文算法的有效性,精心设计了一系列对比实验:(1)单任务学习分类:针对A、B、C这3个地区的样本,利用SVM分类算法分别对各地区样本特征进行分类实验;(2)联合分类:取一定比例的A、B、C这3个地区样本的特征联合作为一个整体样本集,利用SVM分类算法训练出模型后,再分别对各地区剩余样本进行分类实验;(3)多任务学习分类:利用 l_p 范数正则化多任务学习SVM,同时学习A、B、C这3个数据中心样本各自独有的特征和共享特征进行分类。

在实验(1)中,从A、B、C这3个数据中心各选择72例作为训练集,其余作为测试集,将样本特征输入到基于sigmoid核函数的SVM分类算法中,主要参数取值有惩罚因子-*c*为0.05,交叉验证折数-*v*为5,核函数参数系数-*g*为0.05。该算法可通过LIBSVM^[21]实现,并通过交叉验证方法评价分类器的性能,得到分类准确率。在实验(2)中,从A、B、C这3个数据中心各选择72例进行融合,共216例作为训练集,模型训练好后,再对A、B、C数据中心剩余的样本分别进行分类,其他实验条件设置与实验(1)相同。在实验(3)中,从A、B、C这3个数据中心各选择72例作为训练集,其余作为测试集,将样本特征输入到 l_p 范数正则项的多任务学习SVM分类器中进行分类。选择Hinge损失函数作为模型的损失函

数,采用梯度下降法求解式(9)的目标函数优化问题,根据定多变一的原则,得到各参数的最优值。SVM核函数选择高斯核函数,实验中通过交叉验证确定SVM分类器参数 $C=20$, $g=1.2$ 。

2.3 实验结果分析

根据式(5)~式(11),利用梯度下降方法来优化Hinge损失函数,实现 l_p 范数正则化多任务学习SVM分类,并在手写数字数据库(数据集介绍及下载地址为<http://yann.lecun.com/exdb/mnist/>)上实验,验证算法的收敛性如图4所示,随着迭代次数的增加,损失函数值呈不断下降趋势,由此可见,该算法具有较好的收敛性。

各实验获得的最佳分类准确率(ACC)与曲线下面积(AUC)值如表2所示。由表2可知,运用单任务学习算法得到的3个数据中心A、B、C的分类准确率分别达到56.62%、58.97%、60.26%,将各地样本进行融合后,分类准确率分别为60.00%、67.60%、69.00%,而运用多任务学习算法得到的3个数据中心A、B、C的分类准确率分别达到66.67%、75.00%、70.83%,显然较单任务学习分类来说,融合思想对分类结果有一定的提高,多任务学习明显优于单任务学习的分类结果。本实验中多任务学习算法的分类性能之所以优于单任务学习系统,是因为多任务学习过程中考虑了多个任务的关联关系,在同时训练多个任务时,模型利用任务间的共享信息

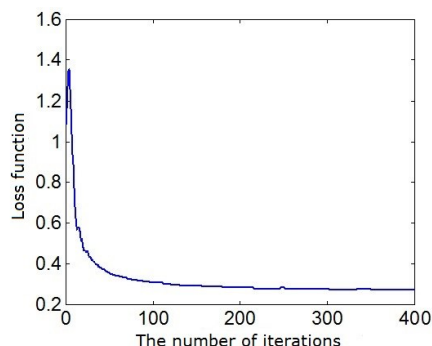


图4 算法收敛性证明

Fig.4 Algorithm convergence

增强系统的归纳偏置能力。同时,加入 l_p 范数正则化项,有效去除了冗余特征,降低了模型的计算复杂度。

表2 对比实验结果(%)

Tab.2 Comparison of experimental results(%)

Experimental method	A		B		C	
	ACC	AUC	ACC	AUC	ACC	AUC
(1)	56.52	0.58	58.97	0.60	60.26	0.56
(2)	60.00	0.65	67.60	0.70	69.00	0.62
(3)	66.67	0.73	75.00	0.72	70.83	0.67

ACC: Classification accuracy; AUC: Area under the curve

3 结语

本研究将图像处理和机器学习方法引入精神分裂症的MRI图像分析中,尝试精确地分类精神分裂症患者和正常人。对于张量形式的MRI图像,提出一种切片化并加权取平均的图像预处理方法,达到降维的效果。其次,对实验样本从统计学角度进行分析,排除性别、年龄因素对实验结果的影响。最后,提取每个人的纹理特征向量,采用单任务学习和正则化多任务学习SVM对正常人和患者进行分类,并对比分析分类结果。实验结果表明,多任务学习明显优于单任务学习,为研究多地区数据和疾病分析提供新思路,更好实现了精神疾病的计算机辅助诊断与预后。

【参考文献】

- [1] 毕文伟. 基于磁共振影像脑疾病分类的模式识别方法研究[D]. 成都: 电子科技大学, 2016: 1-50.
BI W W. A research of pattern recognition methodology for brain diseases & disorders classification with magnetic resonance imaging [D]. Chendu: University of Electronic Science and Technology of China, 2016: 1-50.
- [2] FRIEDMAN L, GLOVER G H, KRENZ D, et al. Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization [J]. Neuroimage, 2006, 32(4): 1656-1668.
- [3] SCHNACK H G, HAREN N E, BROUWER R M, et al. Mapping reliability in multicenter MRI: voxel-based morphometry and cortical thickness [J]. Hum Brain Mapp, 2010, 31(12): 1967-1982.
- [4] BROWN G G, MATHALON D H, STERN H, et al. Multisite reliability of cognitive BOLD data [J]. Neuroimage, 2011, 54(3): 2163-2175.
- [5] SUTTON B, GOH J A, WELSH R, et al. Investigation and validation of intersite fMRI studies using the same imaging hardware [J]. J Magn Reson Imaging, 2010, 28(1): 21-28.
- [6] PEARLSON G. Multisite collaborations and large databases in psychiatric neuroimaging: advantages, problems, and challenges [J]. Schizophrenia bull, 2009, 35(1): 1-2.
- [7] GLOVER G H, MUELLER B A, TURNER J A, et al. Function biomedical informatics research network recommendations for prospective multicenter functional MRI studies [J]. J Magn Reson Imaging, 2012, 36(1): 39-54.
- [8] HORN J D V, TOGA A W. Multi-site neuroimaging trials [J]. Curr Opin Neurol, 2009, 22(4): 370.
- [9] ATTENBERG J, WEINBERGER K, DASGUPTA A, et al. Collaborative email-spam filtering with the hashing trick [C] // Conference on Email and Anti-Spam. 2009.
- [10] 刘成. 基于多任务学习的自然图像分类方法研究[D]. 西安: 西北大学, 2012: 4-52.
LIU C. Research on natural image classification based on multitask learning [D]. Xi'an: Northwest University, 2012: 4-52.
- [11] MARQUAND A F, BRAMMER M, WILLIAMS S C, et al. Bayesian multi-task learning for decoding multi-subject neuroimaging data [J]. Neuroimage, 2014, 92(100): 298-311.
- [12] RAO N, COX C, NOWAK R, et al. Sparse overlapping sets lasso for multitask learning and its application to fMRI analysis [J]. Comput Sci, 2013, 12(5): 2202-2210.
- [13] WATANABE T, KESSLER D, SCOTT C, et al. Multisite disease classification with functional connectomes via multitask structure sparse SVM [C] // Workshop on Sparsity Techniques in Medical Imaging, California, USA: The 11th Conference on Email and Anti-Spam, 2014.
- [14] WANG X, ZHANG T, CHAIM T M, et al. Classification of MRI under the presence of disease heterogeneity using multi-task learning: application to bipolar disorder [M]. Springer International Publishing, 2015: 125-132.
- [15] EVGENIOU T, PONTIL M. Regularized multi-task learning [C] // Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2004: 109-117.
- [16] 刘建伟, 李双成, 罗雄麟. p 范数正则化支持向量机分类算法 [J]. 自动化学报, 2012, 38(1): 76-87.
LIU J W, LI S C, LUO X L. Classification algorithm of support vector machine via p-norm regularization [J]. Acta Automatica Sinica, 2012, 38(1): 76-87.
- [17] SELVARAJ H, THAMARA SELVI S, SELVATHI D, et al. Brain MRI slices classification using least squares support vector machine [J]. Int J Intell Comput Med Sci Image Process, 2007, 1(1): 21-33.
- [18] 李航. 统计学习方法 [M]. 北京: 清华大学出版社, 2012: 10-20.
LI H. Statistical learning method [M]. Beijing: Tsinghua University Press, 2012: 10-20.
- [19] 周志华. 机器学习 [M]. 北京: 清华大学出版社, 2016: 173-177.
ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016: 173-177.
- [20] 浦剑. 多任务学习算法研究 [D]. 上海: 复旦大学, 2013: 47-78.
PU J. On multi-task learning methods [D]. Shanghai: Fudan University, 2013: 47-78.
- [21] 俞斌峰. 多任务学习及其在光谱数据分析中的应用 [D]. 合肥: 中国科学技术大学, 2015: 58-71.
YU B F. Multi-task learning and its application in spectral multivariate calibration [D]. Hefei: University of Science and Technology of China, 2015: 58-71.

(编辑: 陈丽霞)