

基于多尺度眼动信息熵的抑郁检测

王艳¹ 郑秀娟^{1*} 黄申² 苗丹民² 张昀³ 刘凯¹

¹(四川大学电气工程学院 四川 成都 610065)

²(中国人民解放军空军军医大学军事医学心理学系 陕西 西安 710038)

³(西安交通大学电子与信息工程学院 陕西 西安 710049)

摘要 针对传统抑郁检测过程中出现的高主观依赖性和非普适性等弊端,提出一种基于多尺度眼动信息熵的抑郁检测方法。该方法以眼动扫描路径的多尺度信息熵为特征,通过比较被试在同一语义刺激下的眼动扫描路径差异,检测抑郁高危人群和抑郁症患者。在作答抑郁高危自评量表的眼动数据集上,该方法的平均分类准确率为80.36%,相较于MultiMatch、ScanMatch和SubsMatch算法,分别提升了12.50百分点、11.79百分点和9.08百分点。实验结果表明,该方法可以更好地捕获细颗粒度眼动信息,具有更高的准确度和灵敏度。

关键词 眼动扫描路径 相似性分析 信息熵 抑郁

中图分类号 A 文献标志码 A DOI:10.3969/j.issn.1000-386x.2024.10.006

A DEPRESSION DETECTION METHOD BASED ON THE MULTISCALE INFORMATION ENTROPY OF EYE SCANPATH

Wang Yan¹ Zheng Xiujuan^{1*} Huang Shen² Miao Danmin² Zhang Yun³ Liu Kai¹

¹(College of Electrical Engineering, Sichuan University, Chengdu 610065, Sichuan, China)

²(Department of Military Medical Psychology, Air Force Medical University, Xi'an 710038, Shaanxi, China)

³(School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China)

Abstract A depression detection method based on the multiscale information entropy of eye scanpath is proposed to overcome the disadvantages of high subjective dependence and non-universality in the traditional depression detection. Characterized by the multiscale information entropy of the eye scanpath, this method detected high-risk groups of depression and patients with depression by comparing the differences between the subjects' eye scanpaths under the same semantic stimulus. On the eye movement dataset of issues who answer the self-rating high-risk of depression scale, the average classification accuracy of this method is 80.36%, which has 12.50, 11.79, and 9.08 percentage points increase compared with MultiMatch, ScanMatch, and SubsMatch algorithms respectively. Experimental results show that this method can better capture fine-grained eye movement information and has higher accuracy and sensitivity.

Keywords Eye scanpath Similarity evaluation Information entropy Depression

0 引言

抑郁症是影响人类健康最常见的心理疾病之一。其患者常伴随着显著而持久的心境低落、焦虑和运动

性激越,严重者可能会出现幻觉、妄想等精神病症状,严重影响一个人进行日常生活活动的能力^[1]。传统的抑郁高危检测方法依赖于周期性的结构化的访谈和问卷调查,其诊断结果极大程度上取决于心理医师的专业素养与被试的诚实程度,容易受到主观因素的影响,

收稿日期:2021-05-07。国家自然科学基金项目(NSFC61473198);国防科技创新特区项目(18-163-15-ZT-001-007-38);四川省科技计划项目(2019YFS0140);成都市技术创新研发项目(2020-YF05-01386-SN)。王艳,硕士生,主研领域:眼动跟踪技术,数据分析。郑秀娟,副教授。黄申,博士生。苗丹民,教授。张昀,研究员。刘凯,教授。

且耗时耗力不便于展开大规模检测。因此,一个可对抑郁检测进行客观评估且可辅助医师临床诊断的量化指标显得尤为重要。

随着抑郁症的深入研究,人们发现抑郁倾向和认知风格之间存在相关性,可将视觉认知行为模式作为探索抑郁症客观诊断的重要途径^[2]。1971年,Spitz等^[3]提出了眼动扫描路径理论,该理论将眼动扫描路径定义为人类在特定情绪中从特征到特征所执行的特定眼动序列,使人深入了解视觉认知行为过程。1997年,Brandt等^[4]进一步将眼动扫描路径量化为一系列具有时间序列关系的注视点和扫视,并证明了眼动扫描路径在相同刺激下的可重复性。此后,眼动扫描路径广泛应用于视觉信息处理过程中的认知研究,例如场景感知^[5]、阅读^[6]和视觉搜索^[7]。所以,我们可通过比较被试在视觉认知过程中的眼动扫描路径差异来检测抑郁高危人群^[8]和抑郁症患者。

现有的眼动扫描路径相似性分析算法可分为字符串校准、几何矢量、注意力视图和多频模式检测四类。这些算法都是为解决特定问题而开发的,它们侧重于不同维度的眼动扫描路径比较,应用于不同场景中的眼动研究,每种方法都有其优点和缺点。

字符串校准类方法将记录所得的注视点序列映射为字符串,然后根据两字符串对齐后的差异来量化两扫描路径之间的距离。经典算法有编辑距离(Edit Distance)^[9]和 ScanMatch 算法^[10]。在早期探究眼动扫描路径理论的实验中十分流行^[11-12],并且随后广泛应用于各种眼动扫描路径比较工作中^[13-14]。可被用于比较正常和错误回答物理问题的被试之间的眼动扫描路径差异^[15],比较初学者和资深程序员阅读自然语言和源代码的眼动差异^[16]。该类算法简单易用,适用于短且高度相似的扫描路径比较,但不适用于冗长或差异性较大的扫描路径比较,也无法提供具体的眼动特征。

几何向量类算法将眼动扫描路径定义为一系列的几何向量,向量由基于注视点的多维向量及其先前的眼跳构成,将扫描路径对齐问题简化为在表示不同扫描路径的向量之间寻找最小距离映射的问题。该类算法中以 MultiMatch 算法^[17]最受欢迎。可用于探索建筑行业危险识别的视觉搜索策略^[18]。该类算法考虑了时间变化和扫描路径长度,保留了眼动扫描路径的重要特征,但难以确定各向量维度分别适用的特定环境,难以识别扫描路径之间的确切差异。

注意力视图类算法将眼动数据视作一系列没有顺序的点,以这些点为中心的高斯函数叠加形成概率分布图或三维等高线图,然后选择一定的度量方法对上图进

行相似性比较。常见的算法有 iMap 系列算法^[19-21]。可用于探究婴儿对同种族和异族人群面部处理时的眼动扫描路径差异^[22]。该类算法仅测量了特定位置的注视点数量和注视持续时长,虽然可以直观地显示注视点的分布情况,但忽略了时间顺序的重要性,无法识别眼动扫描路径的时序差异。

多频模式检测类算法将扫描路径映射为字符串,通过比较各眼动模式发生频率计算两扫描路径间的相似值。常见算法有 SubsMatch^[23],可用于比较模拟驾驶测试中通过测试与未通过测试的青光眼患者的眼动差异。该类算法避免了标记感兴趣区和时间对齐的问题,可提取高频反复观看眼动模式,常用于动态场景中的眼动扫描路径比较。但该算法时间复杂度较大且参数较多,需反复试验才可获得最优参。

本文提出了一种基于多尺度眼动信息熵的抑郁检测方法。该方法不关注于眼动扫描路径间的局部对齐和比较,而是通过计算正常人、抑郁高危人群和抑郁症患者的眼动扫描路径在不同时间尺度上的信息熵^[24]对各眼动扫描路径的复杂度和不规则度进行整体评估,以分析两扫描路径间的相似性,进而应用分类算法对各类人群进行分类检测。其主要步骤如图 1 所示。

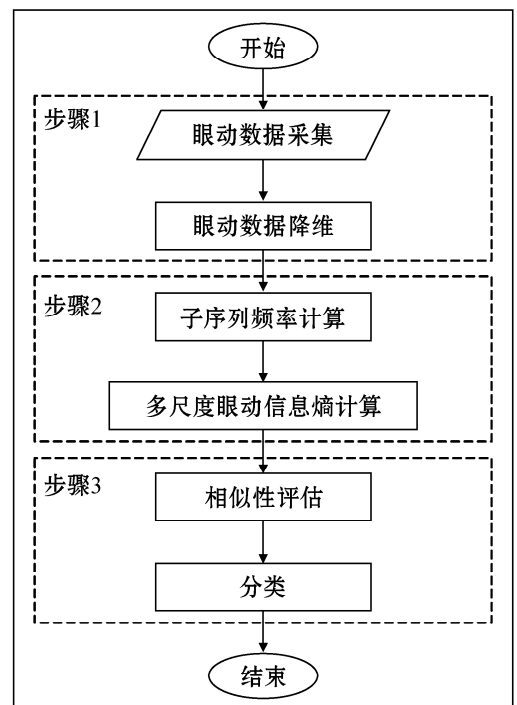


图 1 基于多尺度信息熵的眼动扫描路径相似性分析过程

1 眼动数据降维

眼动扫描路径由一系列具有时间序列关系的注视点和眼跳组成,每个注视点携带其位置和注视持续时

长等信息。为了在保持注视持续时长和时序信息的基础上加快计算速度,本文采用类似符号集合近似 (Symbolic Aggregate AppRoXimation)^[25] 的技术将眼动扫描路径映射为字符串,从而减少了数据维度。其背后的基本思想是在数据上平均分配可用字符,使最终的字符串表达形式中每个字符的个数近似相等。设眼动扫描路径 S 的表达式为:

$$S = \{u_1, u_2, \dots, u_{n-1}\} \quad (1)$$

式中: $u_i (1 \leq i \leq n-1)$ 表示眼动扫描路径 S 中第 i 个眼跳向量,其表达式为:

$$\begin{cases} u_i = f_{i+1} - f_i \\ f_i = (x_i, y_i, t_i) \end{cases} \quad (2)$$

式中: f_i 表示扫描路径 S 中第 i 个注视点, x_i, y_i, t_i 分别表示第 i 个注视点的横、纵坐标和注视持续时长。联立式(1)和式(2),则眼动扫描路径 S 可表示为:

$$S = \{f_1, f_2, \dots, f_n\} \quad (3)$$

设字母表 A 和断点序列 B 为:

$$\begin{cases} A = \{\alpha_1, \alpha_2, \dots, \alpha_a\} \\ B = \{\beta_0, \beta_1, \dots, \beta_{a-1}, \beta_a\} \end{cases} \quad (4)$$

式中: $\alpha_k (1 \leq k \leq a)$ 表示字母表 A 中第 k 个字符; $\beta_j (1 \leq j \leq a-1)$ 表示断点序列 B 中第 j 个断点,其值对应标准高斯分布的随机变量值,且任意两个断点之间对应的标准高斯分布概率值均为 $1/a$, β_0 和 β_a 分别表示 $-\infty$ 和 $+\infty$ 。

将各注视点的横(纵)坐标值做归一化处理,通过式(5)完成注视点到字符的映射。

$$\begin{cases} x_i = \alpha_k & \beta_{k-1} \leq x_i \leq \beta_k \\ f_i = \alpha_k \end{cases} \quad (5)$$

然后,根据注视持续时长将注视点重复映射,使单个字符表示的注视持续时长相等。最终,眼动扫描路径被映射为字符串 $\hat{S} = \{\dots, \alpha_k, \alpha_k, \dots\}$ 。映射过程如图2所示。

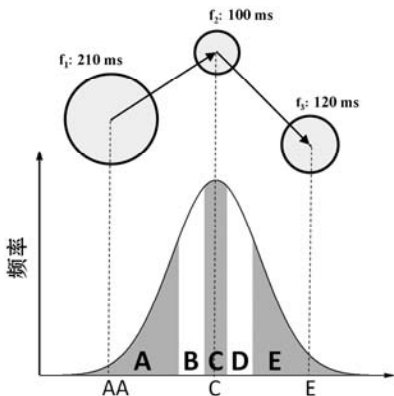


图2 眼动数据字符化映射(时间间隔为 100 ms)

眼动数据字符化的过程不仅约简了眼动数据的表达形式,也在一定范围内消除了眼动扫描路径因设备

校准失误或被试移动等因素造成的空间偏移误差。并且较为完整地保留了注视点坐标、注视持续时长、眼跳、扫描路径长度等关键信息。相较于字符串校准类和多频模式检测类扫描路径比较算法,本文算法在眼动数据降维处理过程中能够更完善地存储眼动信息。

2 多尺度眼动信息熵

多尺度熵是用来表征时间序列复杂程度的重要参数,可在多个时间尺度上求得被测序列的样本熵^[26]。在眼动扫描路径中,不同时间尺度的眼动扫描路径分别表示不同的眼动模式。当时间尺度为 1 个注视点时,此时的眼动模式为注视;当时间尺度为 2 个注视点时,眼动模式为两注视点间的单次眼跳;当时间尺度为 n 个注视点时,眼动模式为 n 个注视点间的连续眼跳。多种眼动模式组成了眼动扫描路径,各眼动模式的信息熵构成了多尺度信息熵的分量,通过比较各眼动扫描路径间的多尺度信息熵差异即可分析其相似性。

在计算眼动扫描路径的多尺度信息熵之前,我们对各子序列(眼动模式)的频率进行定义和计算。采用滑动窗口对字符化后的眼动扫描路径 \hat{S} 进行子序列特征提取,并计算各子序列出现的频率。构建大小为 $w (1 \leq w \leq n)$ 的时间窗口,使其以步长为 1 的速度自序列的头部字符向尾部字符逐步滑动(如图3所示),并顺次累计各窗口中子序列出现的频数。将各子序列频率定义为:

$$P_{\text{subs}} = \frac{n_{\text{subs}}}{l_s - l_{\text{subs}} + 1} \quad (6)$$

式中: n_{subs} 表示子序列在各窗口中出现的频数, l_s 和 l_{subs} 分别为眼动扫描路径 \hat{S} 和子序列长度,即 $l_s = n, l_{\text{subs}} = w$ 。子序列频率与子序列出现频数和子序列长度呈正相关。当子序列长度等于字符串 \hat{S} 的长度或序列 \hat{S} 中仅有一种字符时,子序列频率为 1,当子序列出现频数为 0 时,子序列频率为 0。一般情况下,眼动扫描路径中包含多种眼动模式且眼动扫描路径长度远大于时间窗口大小,不会发生以上两种极端情况。

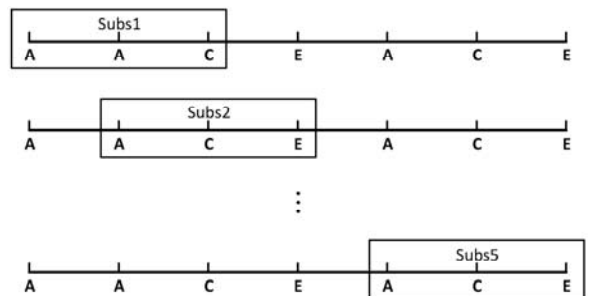


图3 滑动窗口方法($w=3$)

然后,计算眼动扫描路径 \hat{S} 在窗口大小等于 w 时的信息熵。基于信息熵的基本计算式及眼动扫描路径的子序列特征,我们将眼动扫描路径的多尺度信息熵定义为:

$$H = - \frac{1}{\sum_{\text{subs} \in X} n_{\text{subs}}} \sum_{\text{subs} \in X} n_{\text{subs}} p_{\text{subs}} \log_2(p_{\text{subs}}) \quad (7)$$

式中: X 为子序列的集合。由式(7)可知,当子序列高频出现时,即当眼动扫描路中频繁出现相同眼动模式时,该眼动扫描路径的信息熵将减小。而当眼动扫描路径中出现较多不同注视点间的眼跳行为时,其信息熵将增大。因此,眼动扫描路径的信息熵可反映眼动行为的平稳状态。通常,人类在处理视觉信息的过程中将出现大量的注视行为,此时的信息熵较小。

最后,眼动扫描路径在各时间尺度下的信息熵构成了多尺度信息熵的分量,表示为:

$$\vec{S} = \{H_1, H_2, \dots, H_\tau\} \quad (8)$$

式中: $H_w (1 \leq w \leq \tau)$ 表示眼动扫描路径 \hat{S} 在大小为 w 的时间窗口下的信息熵。

3 相似性评估与分类

$S_1 = \{H_{1,1}, H_{1,2}, \dots, H_{1,\tau}\}$ 、 $S_2 = \{H_{2,1}, H_{2,2}, \dots, H_{2,\lambda}\}$ 分别表示眼动扫描路径 S_1 和 S_2 的多尺度信息熵,通过比较向量 S_1 和 S_2 之间的距离评估眼动扫描路径 S_1 和 S_2 的眼动差异。若两向量维度不相等,则需先进行向量对齐再计算两向量间的距离,如式(9)所示。

$$\begin{cases} |H_{2,\mu} - H_{1,1}| = \min_{1 \leq \delta \leq \lambda - \tau} |H_{2,\delta} - H_{1,1}| \\ \text{dist}(S_1, S_2) = \frac{\lambda}{\tau} \sqrt{\sum_{\delta=1}^{\tau} (H_{2,\mu+\delta-1} - H_{1,\delta})^2} \end{cases} \quad (9)$$

式中: $1 \leq \tau \leq \lambda, 1 \leq \mu \leq \lambda - \tau + 1$ 且 $1 \leq \delta \leq \lambda - \tau + 1$ 。一般而言,在眼动扫描路径比较中,时间窗口大小远小于眼动扫描路径长度。过大的滑动窗口无法捕捉高频子序列,使得眼动扫描路径间的信息熵值较大且差异较小,无法进行有效比较。因此,通常情况下 $\tau = \lambda$, 此时 $\text{dist}(S_1, S_2)$ 为向量 S_1 和 S_2 之间的欧氏距离。

以眼动扫描路径 $S_1 = \{ABC\}$ 与眼动扫描路径 $S_2 = \{ABCCAA\}$ 的相似性分析为例。经本文方法计算,眼动扫描路径 S_1 和 S_2 在 $w = 1, 2, 3$ 时的多尺度信息熵为:

$$\begin{cases} S_1 = \{0.5283, 0.5000, 0.0000\} \\ S_3 = \{0.4979, 0.4644, 0.5000\} \end{cases} \quad (10)$$

经式(9)计算,眼动扫描路径 S_1 和 S_2 之间的距离为 0.5022。

基于多尺度信息熵计算两两眼动扫描路径之间的距离,最终得到一个距离矩阵。使用凝聚层次聚类方法或 K-近邻方法等分类器对该距离矩阵进行分类处理,即可获得各被试眼动扫描路径的分类结果,进而识别抑郁高危人群和抑郁症患者。

4 实验与结果

使用 Tobii T60 (60 Hz) 眼动跟踪装置记录了正常人、抑郁高危人群和抑郁症患者在作答抑郁高危自评量表过程中的眼动数据,以评估本文方法对抑郁高危人群和抑郁症患者的检测性能。

4.1 实验对象

正常人:19人,该群体身心健康无精神疾病史且其一级亲属也无精神疾病史,其在一年时间内的多次汉密尔顿抑郁量表^[27]评分均低于7分。

抑郁高危人群:19人,该群体具备抑郁易感人格,并已发展出至少一条抑郁症状却未达到 DSM-5 诊断标准,且社交、职业、生活等有主观不适感,在遇到应激或压力的时候,更容易发展成抑郁症^[3],其一年内的多次诊断结果均达到抑郁高危自评量表筛查标准。

抑郁症患者:18人,该群体符合《精神障碍诊断和统计手册》第五版中抑郁症精神患者的诊断标准,且其汉密尔顿抑郁量表评分均远大于17分。

这三组受试人群在智力和受教育程度上相匹配,无阅读障碍,均可独立完成量表作答。本实验获得了当地伦理委员会的批准,所有被试均签署了实验知情同意书。

4.2 实验数据采集

将抑郁高危自评量表中的59个条目制作成59幅幻灯片作为眼动刺激,并在每个条目前插入了一个校准幻灯片,用于眼动数据采集过程中的坐标校正,如图4所示。每一位被试头部距电脑屏幕60cm,并尽可能保持头部静止。眼动校准后,被试读取幻灯片上的内容,然后单击“是”或“否”选项,以回答问题,每个问题的回答并没有时间限制。单击鼠标即可进入下一张幻灯片。当被试阅读并作答时,我们采用 Tobii T60 (60 Hz) 眼动跟踪仪记录其眼动数据。

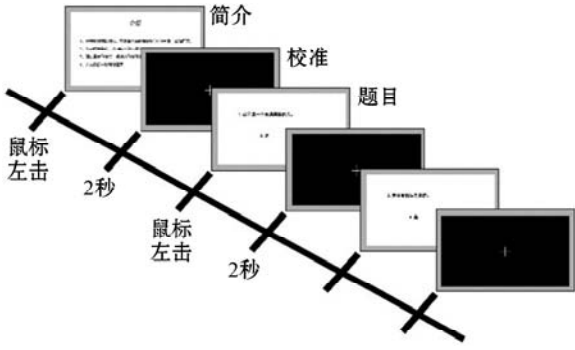


图4 实验流程

4.3 实验结果与分析

时间窗口大小对眼动扫描路径的信息熵计算具有重大影响。图5显示了正常人、抑郁高危人群和抑郁症患者三类人群的眼动扫描路径的信息熵均值随滑动窗口大小而变化(为方便标注,正常人、抑郁高危人群和抑郁症患者分别用标签 C、Hr 和 D 表示)。随着时间窗口大小的增大,信息熵逐渐减小,与此同时,时间窗口大小对信息熵的影响也逐渐减小。这是因为,当时间窗口过大时,无法捕捉高频眼动模式,使得不同窗口大小下的子序列频数和频率近似相等,进而其信息熵变化较小。此外,根据图5显示结果,当滑动窗口大于8时,三类人群的信息熵均值差异甚微,故在本实验中,我们将滑动窗口大小设置为2至8,构建7维多尺度信息熵计算被试间的眼动扫描路径差异。

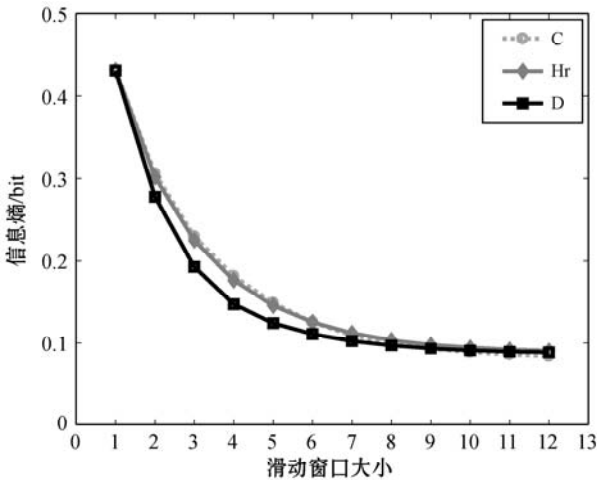


图5 眼动扫描路径在不同时间窗口下的信息熵

选取参数后,基于多尺度信息熵计算两两扫描路径之间的距离,并运用凝聚层次聚类算法对计算结果进行聚类分析,结果如图6所示。抑郁症患者与正常人和抑郁高危人群间的眼动差异较大,被单独聚类在最右边。而正常人和抑郁高危人群间的眼动差异较小,当聚类距离大于1.5时,被聚成一类。尽管正常人和抑郁高危人群的眼动差异较小,但其组间距离仍大于组内距离,当聚类距离为1时,多数正常人被聚为一类,多数抑郁高危人群被聚为另一类。这一结果也与

图中的信息熵均值计算结果相对应,抑郁症患者的信息熵显著低于正常人和抑郁高危人群的信息熵,而正常人与抑郁高危人群的信息熵差异较小,难以区分。

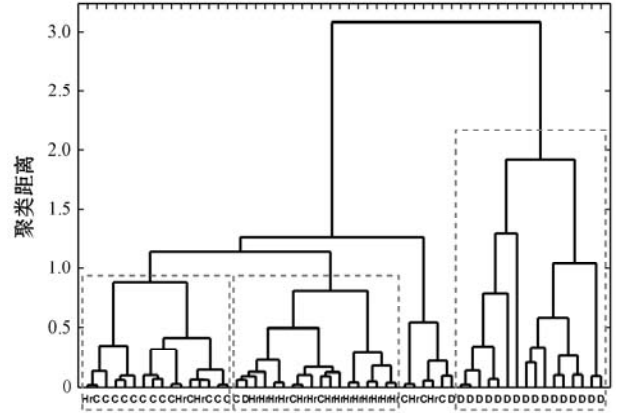


图6 凝聚层次聚类结果

我们将本文方法与现有的 MultiMatch^[17]、ScanMatch^[10] 和 SubsMatch^[23] 算法进行比较。分别运用上述四类算法计算 59 名被试作答抑郁高危自评量表时的眼动差异,并依据计算结果对被试进行分类。平均分类准确度 ACC、宏查准率 macroP、宏召回率 macroR 和宏 F1 分数对上述四类眼动扫描路径相似性分析方法进行性能评估。各评估指标具体计算如下:

$$\begin{cases}
 ACC = \frac{\sum_{i=1}^N TP_i}{TP_1 + TN_1 + FP_1 + FN_1} \\
 macroP = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FP_i} \\
 macroR = \frac{1}{n} \sum_{i=1}^n \frac{TP_i}{TP_i + FN_i} \\
 macroF1 = \frac{2 \times macroP \times macroR}{macroP + macroR}
 \end{cases} \quad (11)$$

式中: n 为标签类别数; TP_i 、 TN_i 、 FP_i 和 FN_i 分别表示第 i 类人群的真正例、真反例、假正例和假反例,详情如表 1 所示。

表1 分类结果混淆矩阵

真实情况	预测结果	
	正例	反例
正例	TP(真正例)	FN(假反例)
反例	FP(假正例)	TN(真反例)

本文方法与 MultiMatch、ScanMatch 和 SubsMatch 算法在本实验数据集上的分类如表 2 所示。在相同分类器的作用下,本文方法分类准确率最高,为 80.36%, F1 系数为 82.36%。相较于 MultiMatch、ScanMatch 和 SubsMatch 算法,本文方法的平均分类精确度分别提升了 12.50 百分点、11.79 百分点和 9.08 百分点,宏 F1

分数最高提升了 10.63 百分点、10.26 百分点和 7.64 百分点。本文方法具有更好的准确性和鲁棒性。

表 2 各算法的分类性能(%)

算法	ACC	macroP	macroR	macroF1
MultiMatch	67.86	76.08	67.86	71.73
ScanMatch	68.57	76.01	68.57	72.10
SubsMatch	71.28	77.90	71.78	74.72
本文方法	80.36	84.07	80.70	82.36

图 7 的混淆矩阵显示了更多分类信息。四类眼动扫描路径相似性分析方法的检测结果均存在显著的假阳性且抑郁患者的检测准确率极高,但抑郁高危人群和正常人较难区分。这也与临床的诊断情况相符,在临床诊断中,被试是否患有抑郁症是较好判断的,但难以通过少次或短时间内的临床检测区分正常人和抑郁高危人群。而本文方法的计算结果优于 MultiMatch、ScanMatch 和 SubsMatch 算法就主要体现在正常人和抑郁高危人群的检测准确度上。这说明本文算法在计算眼动扫描路径间的差异时,具有更高的灵敏度,可以更好地捕获细颗粒度的眼动信息。

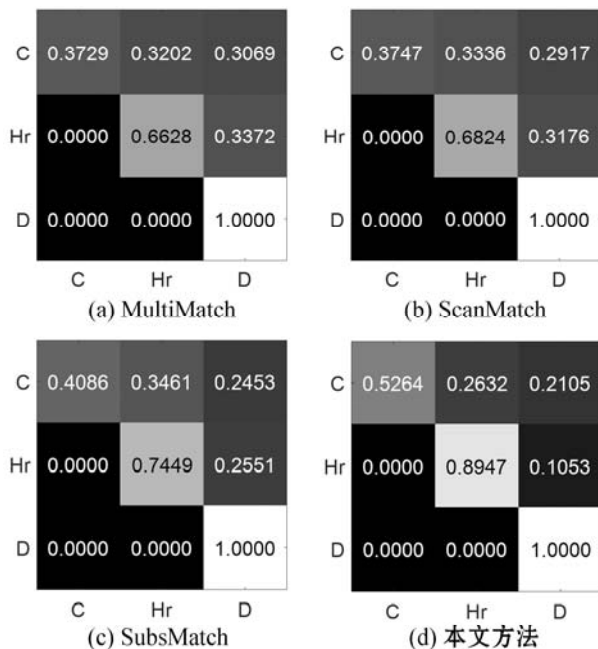


图 7 分类混淆矩阵

5 结 语

本文提出了一种基于多尺度眼动信息熵的抑郁检测方法。该方法将眼动数据映射为一系列具有时空信息的字符串,并通过比较各字符串间的多尺度信息熵计算两两扫描路径之间的距离,然后对距离矩阵进行分类以检测抑郁高危人群和抑郁症患者。在 56 名被

试作答抑郁高危自评量表的眼动数据集上,本文方法可更加精确地识别正常人、抑郁高危人群和抑郁症患者在同一语义刺激下的眼动扫描路径差异,具有更高的准确性和灵敏度。并且该方法在极大程度上解决了传统抑郁症检测过程中诊断主观性以及低效率等弊端,为抑郁症及其高危人群的诊断提供了一种客观且高效的解决方案。

参 考 文 献

- [1] Halfin A. Depression: The benefits of early and appropriate treatment[J]. American Journal of Managed Care, 2007, 13(4): 92-97.
- [2] Bebbington P. Three cognitive theories of depression[J]. Psychological Medicine, 1985, 15(4): 759-769.
- [3] Spitz H. Scanpaths and pattern recognition[J]. Science, 1971, 173(3998): 753-753.
- [4] Brandt S A, Stark L W. Spontaneous eye movements during visual imagery reflect the content of the visual scene[J]. Journal of Cognitive Neuroscience, 1997, 9(1): 27-38.
- [5] Foulsham T, Dewhurst R, Nyström M, et al. Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach[J]. Journal of Eye Movement Research, 2012, 5(4): 1-14.
- [6] Mishra A, Bhattacharyya P. Scanpath complexity: Modeling reading/annotation effort using gaze information: An investigation based on eye-tracking[M]//Cognitively Inspired Natural Language Processing. Berlin: Springer, 2018.
- [7] Dzung R J, Lin C T, Fang Y C. Using eye-tracker to compare search patterns between experienced and novice workers for site hazard identification[J]. Safety Science, 2016, 82: 56-67.
- [8] 常晓,石蕊,辛伟,等.军队开展精神分裂症早期甄别工作设想[J].心理技术与应用, 2014(1): 3-6.
- [9] Marzal A, Vidal E. Computation of normalized edit distance and applications[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 926-932.
- [10] Cristino F, Sebastiaan M, Theeuwes J, et al. ScanMatch: A novel method for comparing fixation sequences[J]. Behavior Research Methods, 2010, 42(3): 692-700.
- [11] Foulsham T, Underwood G. What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition[J]. Journal of Vision, 2008, 8(2): 6-17.
- [12] Underwood G, Foulsham T, Humphrey K. Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition[J]. Visual Cognition, 2009, 17(6/7): 812-834.
- [13] Harding G, Marina B. Real and predicted influence of image manipulations on eye movements during scene recognition

- [J]. *Journal of Vision*,2010,10(2):8-17.
- [14] Foulsham T, Kingstone A. Fixation-dependent memory for natural scenes: An experimental test of scanpath theory[J]. *Journal of Experimental Psychology: General*,2013,142(1):41-56.
- [15] Madsen A, Larson A, Loschky L, et al. Using ScanMatch scores to understand differences in eye movements between correct and incorrect solvers on physics problems[C]//Symposium on Eye-Tracking Research and Applications,2012:193-196.
- [16] Busjahn T, Bednarik R, Begel A, et al. Eye movements in code reading: Relaxing the linear order[C]//23rd International Conference on Program Comprehension,2015:255-265.
- [17] Jarodzka H, Holmqvist K, Marcus N, et al. A vector-based, multidimensional scanpath similarity measure[C]//Symposium on Eye-tracking Research and Applications,2010:211-218.
- [18] Xu Q W, Chong H Y, Liao P C. Exploring eye-tracking searching strategies for construction hazard recognition in a laboratory scene[J]. *Safety Science*,2019,120:824-832.
- [19] Caldara R, Mielliet S. iMap: A novel method for statistical fixation mapping of eye movement data[J]. *Behavior Research Methods*,2011,43(3):864-878.
- [20] McManus C. iMAP and iMAP2 produce erroneous statistical maps of eye-movement differences[J]. *Perception*,2013,42(10):1075-1084
- [21] Mielliet S, Lao J P, Caldara R. An appropriate use of iMap produces correct statistical results: A reply to McManus (2013) "iMap and iMAP2 produce erroneous statistical maps of eye-movement differences"[J]. *Perception*,2014,43(5):451-457.
- [22] Xiao W S, Xiao N G, Quinn P C, et al. Development of face scanning for own-and other-race faces in infancy[J]. *International Journal of Behavioral Development*,2013,37(2):100-105.
- [23] Kübler T C, Kasneci E, Rosenstiel W. SubsMatch: Scanpath similarity in dynamic scenes based on subsequence frequencies[C]//Symposium on Eye-Tracking Research and Applications,2014:319-322.
- [24] Shannon C E. A mathematical theory of communication[J]. *The Bell System Technical Journal*,2001,5(3):379-423.
- [25] Lin J, Keogh E, Wei L, et al. Experiencing SAX: A novel symbolic representation of time series[J]. *Data Mining & Knowledge Discovery*,2007,15(2):107-144.
- [26] Costa M, Goldberger A, Peng C K. Multiscale entropy analysis of physiologic time series[J]. *Physical Review Letters*,2002,89(6):68102.
- [27] Williams J B, A structured interview guide for the Hamilton depression rating scale[J]. *Archives of General Psychiatry*,

1988,45(8):742-747.

(上接第 23 页)

4 结 语

软件测试是软件开发的重要环节,本文分析了 BS10754-1 标准可信系统框架给出的通用原则和通用技术,并以 GB/T 38634.2 定义的 3 层软件测试过程模型为基础,提出了面向可信测试的软件测试过程模型扩充方法,可以帮助企业提升软件测试在验证和确认软件可信方面的可信任能力。未来需要进一步构建详细的测试活动描述和测试方法。

参 考 文 献

- [1] 伊恩·萨默维尔. 软件工程[M]. 10 版. 彭鑫,赵文耘,译. 北京:机械工业出版社,2018.
- [2] 陈汉一,刘春明,韩小波. 在 DO-178C 标准下对军用信息系统软件测试的考虑[J]. *网络安全技术与应用*,2020(10):79-80.
- [3] 陈昕. 可信软件过程质量模型与评估方法研究[D]. 南京:南京航空航天大学,2015.
- [4] British standards institution. 2018 Information technology-systems trustworthiness part 1 Governance and management specification: BS10754-1[S]. London: BSI Standards Publication,2018.
- [5] 陈锦富,卢炎生,谢晓东. 软件错误注入测试技术研究[J]. *软件学报*,2009,20(6):1425-1443.
- [6] 刘凯,梁欣,张俊萍. 软件测试过程模型研究[J]. *计算机科学*,2018,45(S2):518-521.
- [7] 余久久,张佑生. 软件测试改进模型研究进展[J]. *计算机应用与软件*,2012,29(11):201-207.
- [8] 周童,方彦军. 一种新的软件测试过程模型—“跑道”模型[J]. *计算机测量与控制*,2015,23(10):3347-3349.
- [9] 杨阳,吴际,苑春春,等. 面向 DO-178C 软件测试过程的目标符合性论证模式[J]. *计算机研究与发展*,2017,54(3):654-668.
- [10] 文杏梓,罗新星,欧阳军林. 复杂不确定环境下可信软件非功能需求评价模型[J]. *系统工程学报*,2016,31(4):557-567.
- [11] 郭曦. 可信软件测试理论与技术研究[D]. 武汉:武汉大学,2013.
- [12] 陈广勇,祝国邦,范春玲. 《信息安全技术网络安全等级保护测评要求》(GB/T 28448—2019)标准解读[J]. *信息网络安全*,2019(7):1-7.
- [13] 国家标准化管理委员会. 系统与软件工程:软件测试第 2 部分测试过程:GB/T 38634.2—2020[S]. 北京:中国标准出版社,2020.