

生理心理学报告



探究运用不同生理指标测谎的有效性

姓名 : 白张昀凡 鲍研萌 陈仕慨 胡浩宇
刘一诺 温雅岚 薛唯琛 吴晓

实验地点 : 西园 6 舍 114 心理实验室

指导教师 : 王慈老师

探究运用不同生理指标测谎的有效性

摘要 人在提供谎言时会由于自主神经系统兴奋而产生一系列不受意识控制的生理反应,这是生理测谎的理论基础。过往运用多导生理仪进行测谎的研究表明,皮肤电比心电具有更高的准确率,但未从统计角度建立预测模型。本研究利用模拟盗窃情境结合 GKT 测谎模式,验证了皮肤电指标在测谎时具有最高灵敏度。进一步通过支持向量机建模,发现不同测谎标准下模型结果有所区别,但皮肤电或皮肤电结合心电数据进行预测具有稳定的高准确率,而仅利用心电数据无法稳定地有效测谎。

关键词 生理测谎; 多导生理仪; 支持向量机; 皮肤电; 心电

1 研究背景

1.1 生理测谎概述

测谎有着悠久的发展历史。三种常见的测谎方法分别是:测量生理反应、观察行为(尤其是微表情)、分析言语。从 20 世纪 50 年代开始,越来越多的研究者致力于研究言语、非言语和生理线索在测谎研究中的应用,并逐渐发展出各种检测神经活动的非侵入式测谎工具,如脑电图(EEG)、功能性磁共振成像(fMRI)、多导生理仪(polygraph)等(杨 & 傅, 2009)。现代测谎技术最常用的是使用多导生理测试仪(也称多参量心理测试仪)进行测谎,又称自主反应测量技术或生理测谎。

生理测谎的基本假设认为,人在故意提供谎言时会产生一定的心理压力,而这种心理压力又会引起一系列的生理反应,如呼吸速度异常、心跳加快、血压升高、体温微升、手掌出汗、肌肉微颤等。由于这些生理反应都受人的植物神经系统控制,所以它们是人的主观意志所不能制约的(Synnott et al., 2015)。生理测谎技术通过不断完善的仪器测得自主、躯体反应的生理指标,服务于司法犯罪事件测谎、安全部门职员测谎等多个领域(余, 2017)。

测谎的问题主要有三种:准绳问题、相关问题和无关问题。准绳问题的设计用来确定受测人说谎时的生理反应模式。由于每个人在说谎时生理反应模式并不完全相同,所以测谎员在测谎时必须了解每个受测人的具体反应模式,以便作为辨识谎言的标准。

1.2 多导生理仪的应用

测谎过程所激发的情绪是被试害怕谎言被揭穿及对由此而带来的不良结果的心理恐惧,是一种负性情绪。当说谎者试图隐瞒信息的时候,自主神经系统(ANS)会受到不同程度的激活,引起相应生理反应。ANS 由交感神经系统(SNS)和副交感神经系统(PNS)组成,能够通过分泌肾上腺素、去甲肾上腺素、乙酰胆碱等调控人体的心率、血压以及汗腺分泌。多导生理仪测谎主要通过采集皮肤电阻、呼吸、心率血压等自主神经的生理指标来判断被试是否说谎(Fiedler et al., 2002)。

(1) 皮肤电反应

皮肤电反应是人交感神经兴奋性变化的最有效、最敏感的生理参数,是国际上应用最早最广泛并得到普遍认可的测试指标。它通过测量人手心发汗的程度而直接反映出人心理紧张状态的变化。反应幅度大,灵敏度高,不易受大脑皮层意识的直接控制。正是在皮肤电的基础上, Lykken 创造了为全世界广泛采用的 GKT 评分方法(Lykken & Venables, 1971)。

一般情况下,皮肤电流运动具有一定的电阻参数。由于情绪状态下汗腺分泌的变化和皮肤内血管收缩的变化,能引起皮肤电阻的变化。在恐惧状态下,出汗伴随着交感神经的唤起,导致了皮肤电阻下降,皮肤电流增加。皮肤电变化是一种人体生物电现象,我们可以通过仪器记录皮肤电活动。如果给予视、听、痛等感觉刺激及情绪激动时,皮肤电会出现一个大的电位波动,这种由反射导致皮肤上的动作电位波叫做皮肤电反射。Wechsler 等研究了各种生理状态下的皮肤电反射,认为皮肤电变化均与汗腺活动有关,以手掌和脚掌处变化最大,其汗腺功能与身体其它部位之体温调节出汗不同,主要对精神性活动或感受刺激反应最为敏感(Ben-Shakhar, 1991)。因此可以通过多导生理仪捕捉被试的皮肤电活动判断被试的心理活动。

(2) 呼吸反应

呼吸受中央和自主神经系统调节,变化相对多而复杂。实际应用领域的测谎者们经常认为“呼吸”是比“皮肤电”更为灵敏的测谎指标(杜,2009)。但是使用同样的测谎设备进行的实验室实验却表明,皮肤电在敏感性方面明显强于呼吸指标。

人紧张时,呼吸会下意识地发生一系列变化,如深呼吸、屏气、呼吸节律加快或变慢等。当大脑皮层兴奋性发生变化时,必然会影响呼吸波的变化,而这种变化往往是自己意识不到的。当一个特定测试问题威胁到了受测者的安全时,受测者就会产生害怕情绪,这将导致受测者的心理发生变化,心理上的变化将会激活被测试人的交感神经系统,于是呼吸频率加快,并伴有间歇、停顿迹象,幅度变化失去规律性。有经验的测谎专家可以根据自己的经验判断呼吸反应的异常来检测其是否具有犯罪信息。

研究表明,在说谎的情绪冲动下,大约有 30%会出现呼吸短促和加快现象;说谎通常还会造成呼吸速率的减慢,大约有 70%属于此种情形。呼吸变得深长、细而缓慢,速率明显减缓;另外一种相对较少的情形是呼吸速率不规则,时快时慢。

(3) 心血管活动反应

在测谎测试中,对心血管系统生理反应变化的采集也是一个可参考的指标,主要包括心率和血压的变化。在心血管活动调节中,下丘脑是一个非常重要的整合部位,是对机体各内脏机能进行整合的较高级部位。当人在发怒、恐惧等情绪反应时,下丘脑发挥着重要的作用。因此,心理紧张状态和情绪变化,会相应引起心血管活动的改变。事实上,人在心理紧张时,心脏脉搏输出量增加,心律加快,使脉搏波的收缩压上升。另外,人的紧张心理对血液循环的外周阻力带来影响,使外周阻力增加,从而使脉搏波的舒张压上升,这些在图谱上能明显地表现出来(杜亚男,2009)。

1.3 GKT 测谎模式

心理生理学测谎在犯罪调查中的应用和科学研究中有悠久的历史。在司法领域中常用的测试方法目前主要有三种:相关-无关问题测试(R/I),控制问题测试(CQT),犯罪知识测试(GKT)。而在现场调查中,主要运用的是后两种(余,2017)。

GKT 测谎模式的假设是,知情的嫌疑犯在生理上对与犯罪有关的测试项目比与犯罪无关的项目更敏感,在再认犯罪细节知识时会产生加强的生理反应(Lykken, 1981)。问题设置通常为多重选择问题形式,每个问题包括一个相关选项(如,某一个罪行特征)和几个控制选项(内容相似,但却与罪行无关)。罪犯会对相关选项产生更强的防御反应。测试人员一般会依据有关情节的对应率对应相关项目的多少来作出被测人是否知情的结论(唐,2003)。

根据以往 GKT 研究的不同特点, GKT 研究大致可被分成四种范式的研究。其中卡片测试、模拟犯罪测试、个人真实信息测试三种范式为实验室实验研究范式, 真实场景研究范式为现场研究范式(傅, 马 & 丁, 2008)。其中模拟犯罪测试能够通过模拟犯罪行为的实施使被试加深认知加工, 从而使参加实验室实验的犯罪组被试在对犯罪事实的认识上与真实情景更为相似。

1.4 群体的代表性

模拟研究和现场研究的测谎准确率仍然存在不少争议。

部分研究者认为, 由于模拟犯罪是在主试的指导下进行“犯罪”活动, 其犯罪的动机、心理状况与真实情景依然无法接近, 专业人员在实际罪犯审讯中进行的多重检测(即现场研究)才是真正的研究。在实验控制条件下, 由志愿者扮演相应的角色进行模拟犯罪后, 再进行的多重检测(即模拟研究), 与现场研究不同, 其结果也并不十分可信。Carroll(1988)指出, 对于模拟研究中的无辜被试与现场研究中的无辜嫌疑人, 相同的问题对其在情感上有着完全不同的意义, 从无辜者的角度可能会显著地过高估计检测的准确性。因此, 模拟犯罪测试依然只能作为一种探索测谎认知原理的基础性实验方法服务于测谎基础理论的研究。

但同时, McCauley 和 Forman(1993)对相关研究进行了多元分析, 对两种研究类型准确性比较的结论提出了质疑, 表示两种研究类型准确性的差异完全是由模拟研究中的不确定案例造成的。尽管无法完全真实化, 与其他测试方法相比, 模拟犯罪研究具有较一致高准确性、可针对同一“犯罪”活动设计多组相关问题的优势。

因此, 仍然可以采用模拟研究的方式, 对心理生理测谎的不同指标进行有效性验证。

1.5 支持向量机(SVM)

支持向量机(Support Vector Machine, SVM)是一类按监督学习的方式对数据进行二元分类的广义线性分类器, 在有限的数量下性能非常好。

SVM 与其他线性分类器的基本原理相同, 即考虑每个特征的重要性(权重)的前提下把不同的特征综合起来决定标签 y , 并找到一个阈值。当给特征赋权后, 值大于阈值, 分为一种类别; 值小于阈值, 分为另一种类别。在二维空间上, 两类点被一条直线完全分开叫做线性可分。严格而言, 认为两类数据所在集合 D_0 和 D_1 是 n 维欧氏空间中的两个点集, 如果存在 n 维向量 ω 和实数 b , 使得所有属于 D_0 的点 x_i 都有 $\omega \cdot x_i + b > 0$, 而对于所有属于 D_1 的点 x_j 都有 $\omega \cdot x_j + b < 0$, 则认为 D_0 和 D_1 线性可分。要进行分类, 我们需要直线(或平面、超平面)将 D_0 和 D_1 完全正确地划分开, 能实现该功能的 $\bar{\omega} \cdot \bar{x} + b = 0$ 被称作决策边界。

因为我们通过实验得到的数据只是采样, 每个点的实际值都可能在所得数据的周围随机波动, 且每个点的范围大小可能不一样, 因此在选择决策边界时, 需要尽可能让所有点都远离该边界。离决策边界最近的两个(及以上)的点决定了决策边界的宽度, 我们将这些点为支持向量, 而依赖于决策边界宽度的模型就是这里提到的支持向量机 SVM。

我们知道, $\omega \cdot x + b = 0$ 代表的(超)平面两侧乘以非零常数 a 表示的依然是同一个(超)平面, 但 $\omega \cdot x + b = \pm 1$ 与 $a(\omega \cdot x + b) = \pm 1$ 却表示不同的平面。故对一个决策边界, 如果以 $(\omega \cdot x + b) = \pm 1$ 作为宽度的边界, 只需要适当调整 ω 和 b 的倍数, 就能调整决策边界的宽度。

计算决策边界的宽度时, 只需将等式右边化为 ± 1 , 即可算出 $d = \frac{2}{\|\omega\|}$ 。进行分类时, 需要满

足点在决策边界正确的那一侧,且不在决策边界的宽度内。通过凸优化库即可解出该数据集条件下的决策边界。

该分类模型有很多优点:1) SVM 是一种有坚实理论基础的新颖的小样本学习方法。它基本上不涉及概率测度及大数定律等,因此不同于现有的统计方法。从本质上看,它避开了从归纳到演绎的传统过程,实现了高效的从训练样本到预报样本的“转导推理”,大大简化了通常的分类和回归等问题。2) SVM 的最终决策函数只由少数的支持向量所确定,计算的复杂性取决于支持向量的数目,而不是样本空间的维数,这在某种意义上避免了“维数灾难”。3) 少数支持向量决定了最终结果,这不但可以帮助我们抓住关键样本、“剔除”大量冗余样本,而且注定了该方法不但算法简单,而且具有较好的鲁棒性。这种鲁棒性主要体现在增、删非支持向量样本对模型没有影响;有些成功的应用中, SVM 方法对核的选取不敏感。

2 研究目的

本实验在模拟盗窃犯罪情境下对被试进行询问,检测“犯罪”被试在被问及个人盗窃行为时生理反应的变化,以探究多导生理仪测谎的可行性和有效性,考察并比较通过心电反应(RR 间隔、心率)和皮肤电反应两类生理指标进行测谎的准确性;通过有效生理指标建立测谎预测模型,比较模型在本实验条件下与更严格判断条件下的指标稳定性、准确性与贡献度。

3 研究方法

3.1 被试

浙江大学本科生 16 名,年龄 18-21 岁,其中 4 名男性,12 名女性。被试均无心脏问题或其他精神疾病,实验地点为西园 6 舍 114 心理实验室。

3.2 实验设计

本实验为单因素被试内设计。被试内变量为选项类型(问题选项为相关选项/控制选项);因变量为 RR 间隔、心率 HR 和皮肤电活动的振幅。

3.3 实验材料和设备

本实验共需两间实验室,一间为被试接待室,用于向被试讲述实验过程,完成测谎相关信息的获得;一间用于设置犯罪情景,被试在此完成“盗窃”活动;被试完成犯罪后回到被试接待室进行测谎测试。

入室盗窃情境模拟中所需的物品包括行李箱、背包、钱包、人民币、铅笔盒、手机。

测谎部分包括扑克牌测试环节与正式测谎环节,所需物品为扑克牌,利用 Bioradio 便携式多导生理信号采集系统采集被试的心电反应与皮肤电反应。五个电极片分别放置在被试的左、右锁骨下方、掌部大鱼际、小鱼际以及肘部。

3.4 实验步骤

实验包含 2 个阶段(场景图见附录 3):第一阶段为行为模拟阶段,第二阶段为测谎阶段。实验需要 2 名主试配合完成——主试 1 为“引导者”的角色,负责引导被试进行行为模拟;主试 2 为一个“审讯者”的角色,负责对被试进行测谎。

被试来到被试接待室后,主试1按照标准化指导语让被试了解本次实验的内容和过程,并告知将参加一轮独立的测谎测试。然后指导被试进入犯罪情境房间实施盗窃。

盗窃结束后,被试回到被试接待室,连接多导生理仪,由主试2对被试进行测谎。

在正式测谎之前,首先进行“扑克牌测试”:主试2拿出4张扑克牌,让被试选择一张牌拿在手里,自己记住是什么牌,不需要展示,然后要求被试对所有询问的牌都给与否定答案(例如,问:“你抽到的牌是不是梅花3?”答:“不是梅花3。”),告知被试将通过测谎仪去判断他选择的牌。实际上这4张扑克牌的顺序是固定的,根据被试的选择,主试立即就能知道被试选择的是什么牌,接下来依次询问被试手中的牌是否为XX牌,询问结束后告知被试通过测谎仪的结果得知他抽到的牌是哪张牌。“扑克牌测试”的设置是为了确保被试相信“测谎仪能测谎”。

扑克牌测试结束后对被试进行正式测谎。正式测谎包括3组问题,每组包括5个小问题,关键的选项是在先前盗窃任务中的细节。让被试对所有问题选项都在思考后回答“没有”,当被试对关键选项也回答“没有”时,可以推测其正在说谎。具体题目设置见附录2。所有问题都以同样方式提出,尽量减少语言表述上的歧义,相邻问题存在时间间隔,以等待被试生理数据恢复基线水平。在每一组提问中,关键信息总是在第三个小问题被问到,其余四个问题在该组问题中顺序随机。三组提问的顺序在被试间平衡。

测谎结束后,拆除被试身上的器材,请被试回答后测问题,对测试者进行犯罪关键信息的询问,确保对关键信息的掌握。

4 数据分析

4.1 预处理

首先,我们将收集到的心电、皮电数据,使用biocapture软件导出为csv文件。

对心电数据,将cvs文件导入AcqKnowledge并对照原始事件marker(问题回答开始与问题回答结束)打标,再进行0.5~35Hz的带通滤波。接着使用Analysis-Hemodynamics-ECG Interval Extraction计算并导出他们的RR间隔与心率(HR)数据,前者定义为提问结束后10s内心跳周期的平均值,单位为毫秒(ms);后者定义为提问结束后10s内的心跳速度,单位为每分钟跳动次数(BPM)。

对皮电数据,进行0.05-1Hz的带通滤波,再计算每个被试回答每个问题的皮肤电振幅,该指标定义为提问结束后10s内被试皮肤电振幅最高值与最低值之间的差值,需要转化为Z分数。

开始每组提问(碰椅子、偷钱、手机铃声共三组)时,被试的各项生理指标尚不稳定,因此我们舍弃每组提问时第一小问的生理数据,只保留后面四问的数据。我们采集的16名被试中,有1名事先知道了该测谎任务流程,有1名被试的心率变化范围超出3个标准差,1名被试的数据未被完整记录,1名被试实验中出现多次意外,他们的数据被舍弃。

最后我们得到12名被试在回答不同类型问题说谎或未说谎条件下的RR、HR及皮肤电数据,进行下面的分析。

4.2 常规分析

对被试皮电数据进行 2(2 种结果：说谎、未说谎)×3(3 种问题：碰椅子、偷钱、手机铃声)的二因素方差分析发现，问题主效应、问题与结果的交互作用不显著，结果主效应显著 ($F(1,162) = 10.893, p = 0.001$)。对两种不同的结果进行独立样本 t 检验发现，说谎时皮电数据 (Mean = 0.423, SE = 0.146) 显著高于正常回答时 (Mean = -0.127, SE = 0.081), $t = -2.919, p = 0.005$ 。

对被试的心电数据进行相同的方差分析，不论是 RR 还是 HR，结果与问题的主效应都不显著，结果与问题的交互作用也不显著。

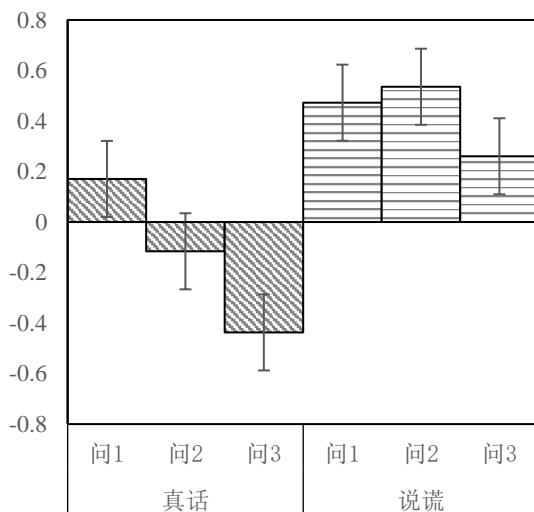


图 1 被试在不同条件下皮电 Z 分数柱形图

4.3 SVM 建模

4.3.1 建模目的

常规分析只能找到对测谎变化最敏感的生理数据，但说谎的生理特征指标个体差异很大，每种特征指标可能反映了不同的心理过程，在说谎与否的条件下变化显著的指标并不一定就能成功预测新被试是否说谎。不过说谎的生理数据一定存在某种共同之处，具有“学习”的可能性，虽然这种模式不容易直接进行定义，但可以通过较大的数据来归纳，故使用生理数据测谎是一个十分适合使用机器学习来解决的问题。

为了从已收集的数据中得到真正有用的信息，我们尝试使用 SVM 来构建预测模型，将从已有数据集上得到的“测谎能力”泛化到未知的数据集上，在下一次测谎过程中，只需要输入被试的生理数据指标，即可用该模型判断其是否说谎。建模过程中我们需要输入一些特征指标，利用不同的特征指标所建立的模型将有不同的预测准确率与稳定性，这能够帮助我们评估哪些指标是有效的。

4.3.2 建模方法

对皮肤电，我们直接用皮肤电变化标准差值与问题序号建立二维向量(x_1, x_2)作为输入值，将问题结果的说谎与否作为输出值 y (定义说谎 $y=1$ ；真话 $y=0$)，匹配得到含有 144 对数据的数据集。对心电图，我们用 RR 值、心率 HR 值及问题序号建立三维向量(x_1, x_2, x_3)作为输入值；对心电+皮电组，我们用皮电变化标准差值、RR 值、心率 HR 值及问题序号组成输入

的四维向量(x_1, x_2, x_3, x_4), 输出值与皮电组相同。最终得到 3 组各 144 对数据的数据集。

每组数据, 用随机种子将数据集随机分为 30% 的测试数据与 70% 的训练数据; 再利用 SVM 算法(见 1.5)在训练集上计算参数 ω 和 b , 根据训练集的正确率调整超参数, 使得训练集的分类准确率尽可能的大(但避免过拟合); 最后, 使用测试集数据验证该决策边界下判断的准确率。

由于提升数据量可以提高训练结果的泛化能力, 为充分利用数据, 我们对原数据进行了 100 次的按比例随机拆分。划分时, 从 0~99 给随机种子定义了不同的参数, 确保每次生成的随机数序列不同, 得到不同特征输入下预测的准确率变化(图 2)。

4.3.3 本实验条件所得建模结果

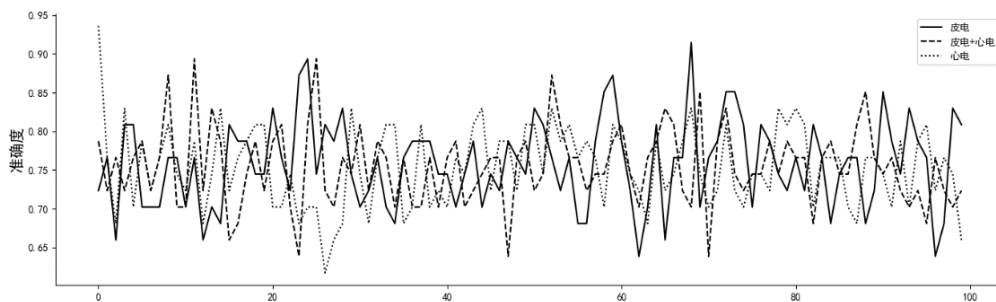


图 2 不同随机种子参数条件下不同成分所建 SVM 模型预测准确率变化图

皮肤电 ($t = 12.445, p < 0.001$)、皮肤电结合心电 ($t = 11.177, p < 0.001$) 及心电 ($t = 12.650, p < 0.001$) 作为输入的组别, 其最大准确率都显著高于随机水平 (75%)。其最大准确率数据见图 3 和表 1。

表 1 不同特征输入下模型预测的准确性

指标	Mean	SE
皮电	0.9149	0.0510
心电	0.9362	0.0518
皮电+心电	0.8936	0.0538

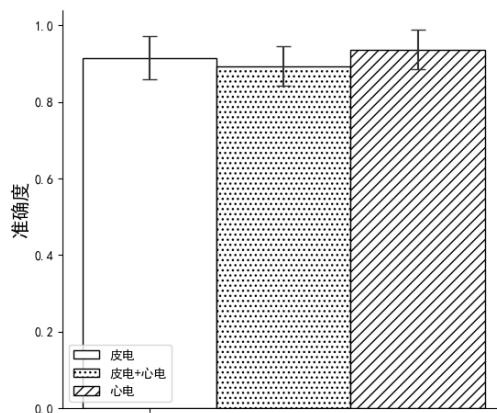


图 3 不同特征输入下模型预测的最大准确率

皮肤电与心电结合进行建模分析, 可以得到在最大预测准确率情况下的各成分贡献度,

具体见图 4。为该模型提供最大贡献度的成分是心电 RR 间隔，其次是心电 HR，皮电数据也有一定程度的贡献。其中，负贡献度表示建模过程中需要对数据取相反数。

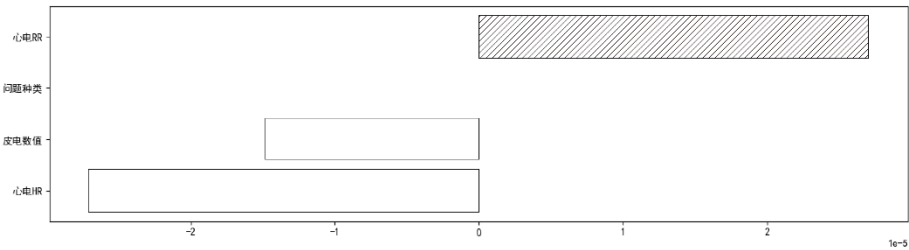


图 4 综合预测模型下各组分的贡献度

使用五折交叉验证法确定这些模型的稳定性，具体操作流程如图五。将训练集随机平均分为 5 份，依次选择一种一份作为验证集，剩余的仍然为训练集（称为“某一折”）即最后有 5 折。对这 5 个折，分别通过训练集算出参数，然后在验证集上算出准确率，通过比较这些准确率之间的差异，我们就能了解模型的稳定性。对各种指标所训练出的模型取平均，由图 6 可见，各组稳定性相同。

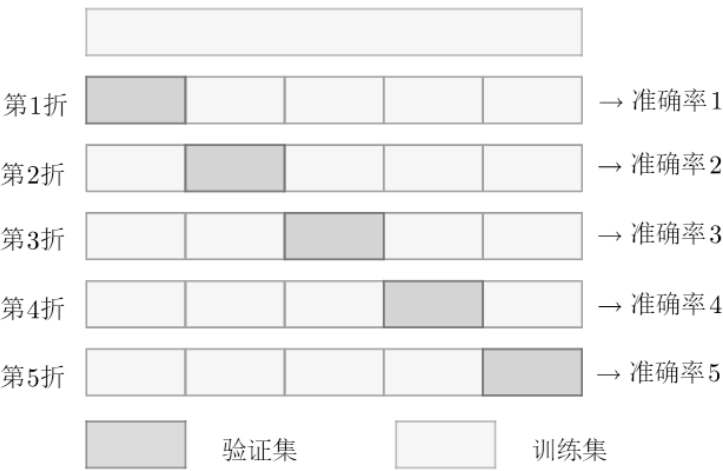


图 5 五折交叉验证法流程 (k=5)

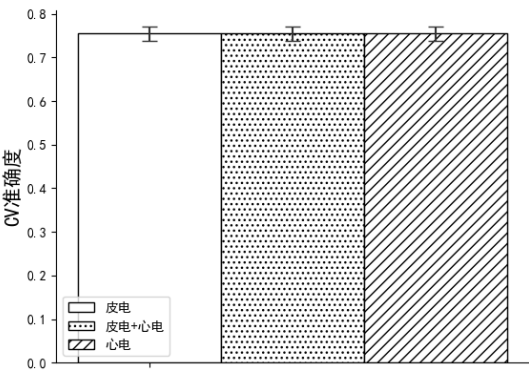


图 6 五折交叉验证法所得稳定性

4.3.4 严格测谎条件下建模结果

在本研究的实验设计下，每组问题共五小问，为等待指标稳定而除去第一小问后，每组问题仍有四小问，被试仅在其中一小问中说谎，即本实验条件下说谎概率仅为 25%。但在实际应用中，总体而言，被试在一个问题上说谎的概率为 50%。为了更贴近现实应用，我们在模型中进行加权，使对说谎的检测标准更为严格。基于此所得模型结果与上述结果有所区别。

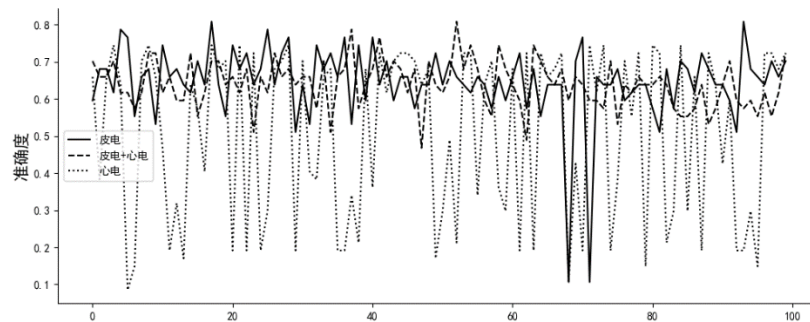


图 7 不同随机种子参数条件下不同成分所建 SVM 模型预测准确率变化图

皮肤电 ($t = 13.827, p < 0.001$)、皮肤电结合心电 ($t = 13.299, p < 0.001$) 及心电 ($t = 9.241, p < 0.001$) 作为输入的组别，其最大准确率仍然显著高于随机水平 (50%)，具体数值见图 8 和表 2。且稳定性较好，两组间不存在显著差异。但仅以心电作为输入时模型预测结果并不稳定，准确率波动大，多数低于随机水平，总体准确率最低。

表 2 加权后不同特征输入下模型预测的准确性

指标	Mean	SE
皮电	0.8298	0.0826
心电	0.7447	0.2180
皮电+心电	0.8298	0.0859

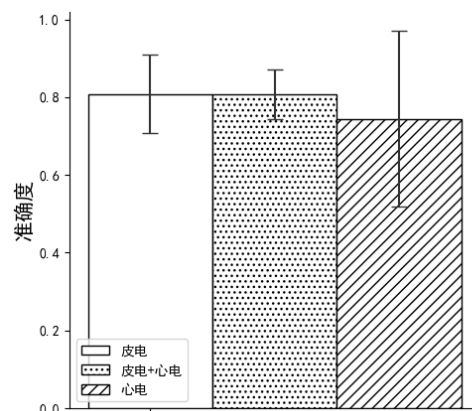


图 8 加权后不同特征输入下模型预测的最大准确性

最大预测准确率情况下的各成分贡献度与前述模型结果区别较大。由图 4，皮肤电数据对于测谎具有极高的贡献度，问题种类和心率也具有一定的贡献度；心电 RR 间隔有一定的

负贡献度。各成分的整体贡献度正负分布与前一模型的结果相反。

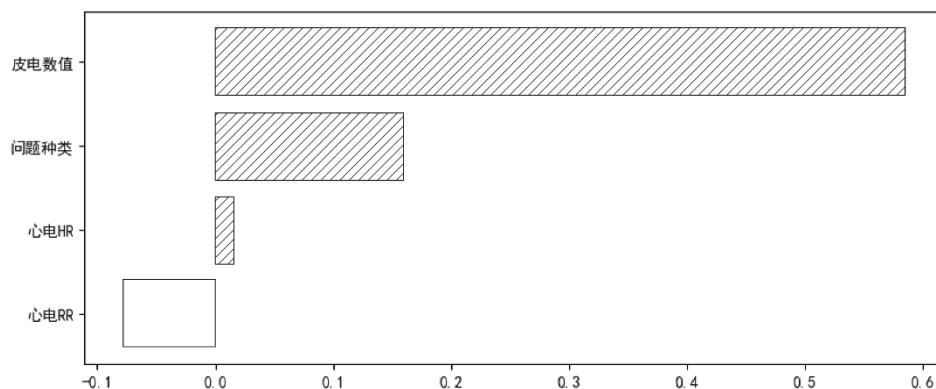


图 9 综合预测模型下各组分的贡献度

同样使用 5 折交叉验证法确定这些模型的稳定性。由图 6 可见，心电组稳定性最差。

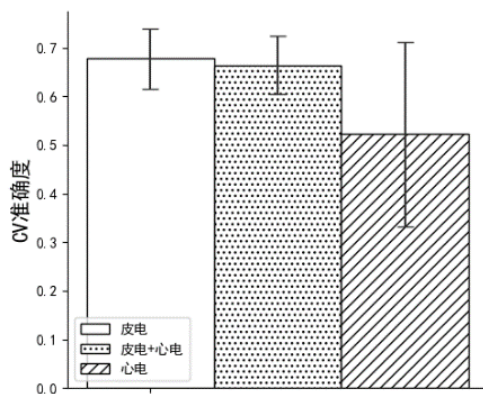


图 10 五折交叉验证法所得稳定性

5 讨论

说谎情况下，被试的皮肤电幅值显著上升，而心电的 RR 和心率 HR 变化均不显著。这可能是因为被试说谎时情绪紧张激动，且手掌处的汗腺活动较为敏感，受到这种精神性活动的调制，导致皮肤上的动作电位产生较大幅度的波动。虽然心理紧张状态下心脏脉搏输出增加，心率加快，可以体现呼吸变化的 RR 值也应有所下降，但该实验条件下，这些数据的变化没有表现出统计显著性，可能由于模拟盗窃下被试的紧张程度未能达到足够水平，与此前研究中皮肤电比心电指标更为灵敏的研究结果基本一致。

在本实验条件下进行 SVM 建模后我们发现，使用心电、皮肤电或心电结合皮肤电进行预测的准确率都显著高于平均水平，测谎有效。其中，使用心电三维向量的最大预测效果最好，心电结合皮肤电的最大预测效果最差。这说明被试在说谎时心电确实发生了变化，只是个体差异太大，不能仅从平均值中发现显著变化，但利用这些数据训练出的模型稳定，提取到的有效信息量大。且被试在说谎过程中皮电的变化模式与心电可能相左，导致增加特征量的输入后模型稳定性反而下降，信息熵增加，预测效果变差。当然，心电数据预测准确性好可能是因为我们预处理分解出了记录心电中更多更细化的组分，但对皮肤电只进行了简单的

标准化处理,即数据预处理的精细度可能也对模型的预测准确性造成影响。

为了进一步将模型应用于实际,我们采用了更严格的测谎标准,建立新的 SVM 模型,得到了不同的结果。在严格判断条件的模型下,使用皮肤电或心电结合皮肤电进行预测的准确率都显著高于平均水平,稳定性较好,测谎有效;而仅凭心电进行预测的准确率较低,稳定性差。同时,贡献成分分析中表明皮肤电指标具有极高正贡献值。这些结果都与最初方差分析的结果一致,共同说明心电指标在测谎中的灵敏性不够,而皮肤电指标具有较好的测谎能力。贡献成分分析中还发现问题种类具有较高正贡献值,可能是由于不同问题在整个盗窃过程中的关键程度不同。

两个模型的贡献成分分析中,各指标具有的贡献情况相反,我们的理解是,基于本实验设计建立的模型所预测的是“说真话”的概率,而基于严格测谎判断下建立的模型预测的是被试“说谎”的概率,因此所有指标呈现相反的贡献值。同时这也说明,心电 RR 在预测“说真话”时具有较好的能力,但要预测“说谎”,仍然是皮肤电指标最为有效。两个模型之间的区别说明测谎的判断严格程度会部分影响测谎结果,但皮肤电是在模型间稳定、有效的指标,结合皮肤电和心电也同样行之有效,但无法仅凭心电测谎。

该实验中我们所用被试量较少,但已得出了较为显著的结果,因此在增加被试量的情况下测谎效果将会更好。事实上,机器学习还有很多可用算法,SVM 算法并不一定是该情况下的最优解,故在后续研究中,或许可以尝试其他建模方法,比较对相同的生理指标以不同算法进行建模预测性的有效性区别。

6 结论

该研究证明了使用多导生理仪器记录生理数据进行测谎的可行性。仅通过方差分析,提问结束后 10s 内被试皮肤电振幅最低值与最高值之间的差值对测谎敏感,而心电相关的数据并不太敏感。使用支持向量机建立预测模型,模型结果受测谎判断严格性影响。在本实验条件下,利用在测谎场景下收集到的心电和皮肤电数据可构建出高准确率的测谎模型,仅利用心电数据时所构建的模型有最高的最大预测准确度。而在实际应用的严格测谎标准下,皮肤电或皮肤电结合心电仍然是稳定有效的指标,仅心电指标无法有效测谎。

参 考 文 献

- [1]杜亚男.(2009). GMT 测谎模式的有效性及其改进[硕士学位论文,浙江师范大学].
- [2]傅根跃;马艳;丁晓攀;(2008). GKT 测谎模式研究. 中国临床心理学杂志, 01, 106 - 109.
- [3]唐兢.(2003). 论犯罪知情测试相关项目的有效性. 政法学刊, 04, 61 - 63.
- [4]杨越, & 傅根跃.(2009). 测谎研究的新进展. 应用心理学, 15(04), 322 - 328.
- [5]余润鑫.(2017). 模拟海关情境下测谎的生理机制初探[硕士,浙江师范大学].
- [6]张钦廷, 黄富银, 洪武, 方明成.(2005). 多道心理生理测谎的心理测量学要素. 中国司法鉴定, 6: 39-41.
- [7]Ben-Shakhar, G. (1991). Clinical judgment and decision-making in CQT-polygraphy. *Integrative Physiological and Behavioral Science*, 26(3), 232-240.
- [8]Fiedler, K., Schmid, J., & Stahl, T. (2002). What Is the Current Truth About Polygraph Lie Detection? *Basic and Applied Social Psychology*, 24(4), 313-324.

[9]Lykken, D. T., & Venables, P. H. (1971). Direct measurement of skin conductance: A proposal for standardization. *Psychophysiology*, 8(5), 656–672.

[10]Synnott, J., Dietzel, D., & Ioannou, M. (2015). A review of the polygraph: History, methodology and current status. *Crime Psychology Review*, 1(1), 59–83.

附录

1 犯罪模拟情境设置

要求被试想象自己是一名小偷，进入房间盗取装在行李箱中的人民币。被试推开门进入房间门没锁，门前有一把抵着门的椅子。房间里地上有一个行李箱。旁边桌上贴有指导语（“请根据铅笔盒里的密码打开旁边的行李箱，盗走钱包中的人民币，请将盗走的钱装进自己的口袋后再离开实验室”），还有一个铅笔盒。小偷在铅笔盒内发现写着密码的小纸条，根据纸条上的密码打开行李箱。于是发现一个背包，背包里有一个手机和一个钱包。被试打开箱子的同时，手机铃声会响起。被试打开钱包偷走里面的一百元钱装进口袋里完成实验。

根据此情景，拟定三组关键信息：

1. 进门时推开了一把椅子。
2. 偷走了一百元钱。
3. 打开箱子的同时，手机铃声响起。

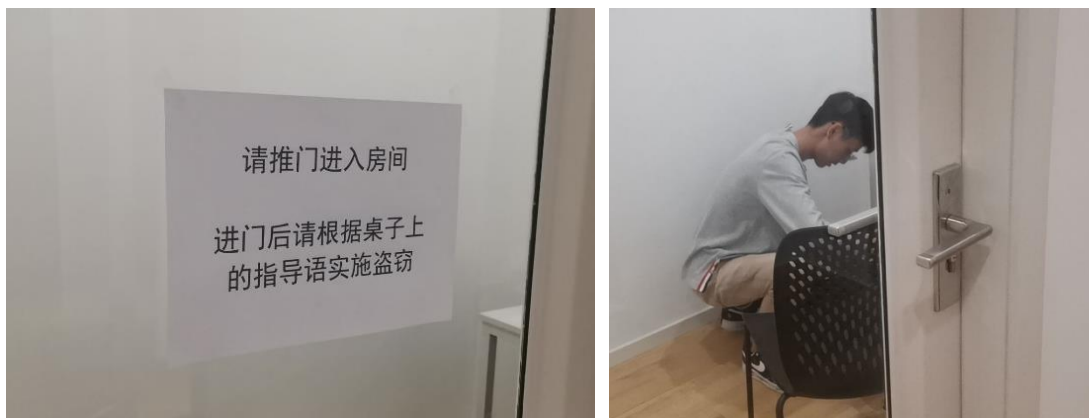
2 GKT 问题

对于以下问题，被试均需回答“没有”。

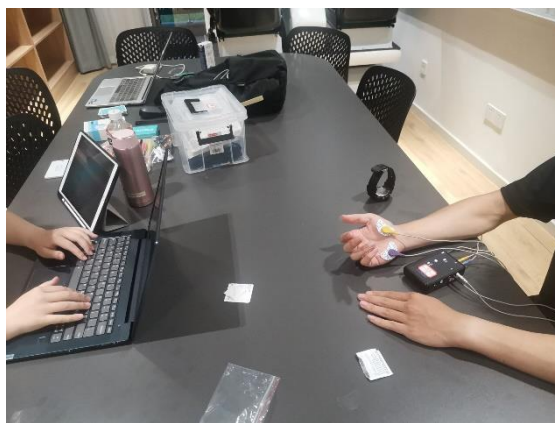
- (1) 小偷进入实验室时，门口有_____吗？
 - A. 一个桌子
 - B. 一个柜子
 - C. 一把椅子
 - D. 一把雨伞
 - E. 一个箱子
- (2) 小偷离开实验室时，带走了_____吗？
 - A. 钥匙
 - B. 手机
 - C. 钱
 - D. 手表
 - E. 钱包
- (3) 小偷打开行李箱时，听到_____的声音了吗？
 - A. 说话
 - B. 脚步
 - C. 手机响了
 - D. 鸟叫
 - E. 敲门

3 实验场景图

阶段 1



阶段 2



Explore the effectiveness of using different physiological indicators to detect polygraphs

Abstract When people lie, they will produce a series of physiological responses that are not controlled by the consciousness due to the excitement of the autonomic nervous system. This is the theoretical basis of physiological lie detection. Previous studies on polygraphs using multi-lead physiometers have shown that the skin electrogram has a higher accuracy rate than the electrocardiogram, but no predictive model has been established from a statistical perspective. This study uses simulated theft scenarios combined with GKT polygraph mode to verify that the electroskin index has the highest sensitivity in polygraph detection. Further through support vector machine modeling, it is found that the model results are different under different polygraph standards, but the skin electricity or the skin electricity combined with the ECG data for prediction has a stable and high accuracy rate, and the ECG data alone cannot be used for stable and effective measurement of lie.

Key words physiological polygraph; multi-channel physiometer; support vector machine; electroskinogram; electrocardiogram