VR-LENS：基于超学习的晕动症检测与可解释人工智能引导的虚拟现实部署

RIPAN KUMAR KUNDU,美国密苏里大学哥伦比亚分校
OSAMA YAHIA ELSAID,美国密苏里大学哥伦比亚分校
PRASAD CALYAM,美国密苏里大学哥伦比亚分校KHAZA
ANUARUL HOQUE,美国密苏里大学哥伦比亚分校

虚拟现实（VR）系统容易引发晕动症，这会严重影响用户体验。因此，近期大量研究提出了多种基于机器学习（ML）和深度学习（DL）的自动化方法来检测晕动症。然而，这些检测方法被认为计算开销大，且属于黑箱方法。因此，这些技术既不可信，也不适用于部署在独立式虚拟现实头戴显示器（HMD）上。本研究提出一种基于可解释人工智能（XAI）的框架 *VR-LENS*，用于开发晕动症检测机器学习模型、解释模型、缩减模型规模，并将其部署到搭载高通骁龙750G处理器的三星A52设备上。具体而言，我们首先开发一种新颖的基于超学习的集成机器学习模型，用于晕动症检测。接下来，我们采用事后解释方法（如SHAPley加性解释（SHAP）、Morris 敏感性分析（MSA）、局部可解释模型无关解释（LIME）和部分依赖图（PDP））来解释预期结果并识别最具主导性的特征。随后，使用已识别的主导特征重新训练晕动症超级学习器模型。对于集成传感器、游戏玩法和生物生理数据集，我们提出的方法识别出眼动追踪、玩家位置以及皮肤电/心率反应是最具主导性的特征。我们还证明，所提出的可解释人工智能引导特征降维方法在保持基准准确率的同时，将模型训练时间和推理时间分别显著缩短了1.91倍和2.15倍。例如，使用集成传感器数据集时，我们的精简超级学习器模型优于最先进研究：它将晕动症分为4类（无、低、中和高），准确率达到96%；并对快速晕动症量表（FMS）1–10进行回归，均方根误差（RMSE）为0.03。我们提出的方法可帮助研究人员实时分析、检测并缓解晕动症，并将基于超级学习器的晕动症检测模型部署在一体式虚拟现实头显上。

CCS 概念： • **以人为中心的计算**！ **人机交互（HCI）**； 虚拟现实；网络晕动症； • **以人为中心的计算**！设计与评估方法。

其他关键词和短语：虚拟现实、晕动症检测、可解释人工智能、机器学习、降维、模型部署

ACM 引用格式：
Ripan Kumar Kundu、Osama Yahia Elsaid、Prasad Calyam 和 Khaza Anuarul Hoque。2023 年。VR-LENS：基于超学习的晕动症检测与可解释人工智能引导的虚拟现实部署。载于 。ACM，美国纽约州纽约市，24 页。
https://doi.org/XXXXXXX.XXXXXXX

特此免费授予许可，允许为个人或课堂使用而制作本作品全部或部分内容的数字或纸质副本，前提是不得为营利或商业利益而制作或分发副本，且副本须附有本声明及首页完整引文。对于本作品中由 ACM 以外的其他方拥有版权的组件，必须尊重其版权。注明出处后，允许摘录。若要以其他方式复制、重新出版、发布到服务器或重新分发至列表，须事先获得特定许可和/或支付费用。请向 permissions@acm.org 申请许可。© 2023 计算机协会。稿件提交至 ACM

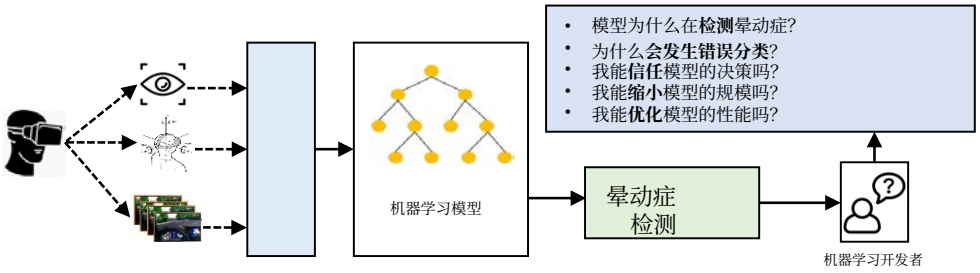


图 1. 基于机器学习的晕动症（网络晕动症）检测中的关键问题。

1 引言

虚拟现实（VR）正应用于包括国防 [2], 教育 [59], 医疗保健 [27],公共安全 [53], 等在内的各种 fi领域 [7, 19, 70, 76]。具体而言，在当前 COVID-19 疫情形势下，VR 为远程学习 [63], 、作为现实协作工作空间 [88, 98], 以及作为成年人的心理健康应对策略 [19] 提供了巨大前景，因为它能够提供人的临场感。然而，VR 用户经常会出现虚拟现实眩晕或晕动症，从而妨碍其沉浸式体验。因此，晕动症已成为阻碍 VR 更广泛接受的重要因素 [59, 76, 87] 。晕动症可定义为一系列令人不适的症状，例如眼疲劳、头痛、恶心、定向障碍，甚至呕吐 [22, 50, 70, 87]。检测晕动症的一种常用技术，是使用沉浸体验后的主观问卷，例如模拟器晕动症问卷（SSQ）和虚拟现实晕动症问卷（VRSQ） [84]。相比之下，快速晕动症量表（FMS） [43] 可以在沉浸过程中评估晕动症严重程度。然而，FMS 依赖用户反馈，即需要在 VR 沉浸过程中进行人工干预 [62]。为克服这些局限，深度学习（DL）和机器学习（ML）近年来已成为晕动症检测的热门方法 [1, 31, 32, 38, 41, 50, 75]。

先进的机器学习/深度学习模型能够根据生理信号（例如心率（HR）、皮肤电反应（GSR）、呼吸频率（BR）、脑电图（EEG）） [19, 65, 75, 91, 100],集成的眼部和头部追踪传感器 [33], 以及立体视频 [33, 52, 67], 以较高的准确率检测晕动症。例如，Islam 等人 [33]提出了一种基于深度融合网络的晕动症严重程度检测方法, 该方法利用用户的眼动追踪和头部追踪数据，准确率达到 87.7%。其他研究人员使用脑电图（EEG） [75, 91, 100], 立体视频 [33, 67]和生物生理信号 [31, 75]数据，以较高的准确率检测晕动症。尽管机器学习/深度学习模型在晕动症检测方面前景广阔，但这些方法仍存在若干缺点（见图 1）。

- 目前最先进的 VR 晕动症检测机器学习/深度学习模型大多依赖黑箱模型，因此缺乏可解释性。为这些黑箱模型增添可解释性，可以显著fi提升模型的可信度，并深入了解机器学习/深度学习模型为何以及如何得出特定fi决策。
- VR传感器会生成海量数据，导致模型复杂、规模庞大且耗电量高，因此在独立式头戴设备（HMD）上进行实时晕动症检测颇具挑战。为机器学习/深度学习模型添加解释，可以帮助工程师通过识别主导特征，有效优化模型。事实上，降低晕动症机器学习/深度学习模型的维度，可以显著缩短训练时间和推理时间，并减小模型大小，同时不影响准确率。

为说明所提出的可解释人工智能方法的必要性，不妨设想一位 VR 游戏开发者希望基于使用资源受限的独立式 VR 头显的 Alpha/Beta 玩家，为其游戏开发晕动症检测模型。假设开发者在游戏中使用了一个缺乏可解释性的黑盒晕动症检测模型，那么他们就难以确定模型的哪些特征（例如来自世界空间的注视原点）会影响晕动症预测。因此，他们只能通过反复试验来缩小模型大小，以尽量减少资源消耗，因为 Meta Quest Pro 的资源本就有限。

值得一提的是，将可解释人工智能（XAI）应用于晕动症解释方面，仍存在显著的研究空白。最近，Kundu 等人 [49] 使用内在可解释的机器学习模型，基于生理[35] 和游戏数据集[73]进行晕动症检测与解释。然而，他们的研究采用的是二元分类器，只能检测是否存在晕动症。这类简单模型无法检测晕动症严重程度，因此不适用于真实的VR应用。此外，内在可解释模型通常依赖数据属性，因此容易受到维数灾难问题的影响 [83]。例如，基于决策树和逻辑回归的内在可解释模型可能会因维度和非线性问题而出现过fi拟合，最终导致其在晕动症分类中的性能不佳 [49, 69, 75]。大多数先前研究使用特征选择或降维技术（例如主成分分析（PCA） [45, 47, 60, 82, 89] ），以解决晕动症检测机器学习/深度学习模型的高维问题。例如，Lin 等人 [55] 使用结合PCA的机器学习模型提取与晕动症相关的特征，以预测晕动症等级。同样，Kottaimalai 等人 [48] 使用PCA降低脑电信号的维度、复杂度和计算时间，从而检测晕动症。然而，应用PCA会导致信息丢失 [3]。此外，Mawalid 等人 [61]使用时域特征提取方法提取脑电统计特征，以对晕动症进行分类。然而，应用PCA会导致信息丢失 [3]。这意味着，基于PCA的降维可能会丢失对准确检测晕动症至关重要的重要特征。

为应对上述挑战，我们提出了一种新颖的方法 VR-LENS——一个基于可解释人工智能的框架，用于晕动症检测、解释、模型缩减和模型部署。首先，为展示所提方法的适用性，我们提出了一种新颖的基于超学习的集成机器学习模型。随后，我们采用事后解释方法，包括 SHapley 加性解释（SHAP）、Morris 敏感性分析（MSA）、局部可解释模型无关解释（LIME）和部分依赖图（PDP），来解释预期结果并确定最主要的特征。具体而言，我们首先开发了一种新颖的基于超学习的晕动症检测集成机器学习模型。接着，我们采用事后解释方法，即 SHapley 加性解释（SHAP） [57], Morris 敏感性分析（MSA） [64], 局部可解释模型无关解释（LIME） [80],和部分依赖图（PDP） [24], , 为分析、识别和排列导致晕动症的主导特征提供全局和局部解释。随后，我们使用识别出的主导特征重新训练超级学习器模型，即使用较少的特征进行训练。由此得到的轻量级晕动症检测模型所用特征数量显著减少。最后，为展示 VR-LENS 框架的有效性，我们将精简后的超级学习器模型部署在基于高通骁龙750G处理器的三星 A52 设备上 [30]，因为许多最先进的 VR 设备均采用高通骁龙处理器。结果表明，所提出并部署的精简超级学习器模型在部署设备上训练速度更快、推理时间显著缩短，同时准确率优异，表现优于最先进研究。例如，我们的结果表明，使用集成传感器数据集 [33],时，采用全部特征的所提超级学习器模型能够以 95% 的准确率将晕动症严重程度分为 4 类（无、低、中和高），并以 0.04 的均方根误差预测正在发生的晕动症（FMS 1–10）。

（RMSE）为 0.04。然而，对同一超级学习器模型和同一数据集进行基于可解释人工智能的缩减后，仅使用基线模型三分之一的特征，我们就能以 96% 的准确率对晕动症严重程度进行分类，并以 0.03 的 RMSE 预测正在发生的晕动症。此外，我们的缩减方法使部署在嵌入式设备上的训练时间提升了 1.91 倍，推理时间提升了 2.15 倍。据我们所知，这是首项使用超级学习器模型和事后解释技术来进行晕动症检测、解释、模型缩减和部署的研究。因此，我们认为，所提方法有助于未来研究人员理解、分析并设计更简洁、轻量、可信且准确的晕动症检测模型，使其适用于独立式头戴显示器上的实时实现。

2 相关工作

“晕动症 ”一词指一系列症状，例如头晕、恶心等，这些症状与晕动病相似，可能发生在沉浸式体验期间或之后 [51]。解释晕动症成因的最广为接受的理论是感官冲突理论fl [50, 51]。该理论认为，当眼睛感知到运动，而前庭系统却没有感知到时，就会发生晕动症。不过，中毒理论和姿势不稳定理论等其他理论也被用于解释晕动症的成因。此外，用户的年龄、性别和此前的虚拟现实体验等因素也会影响晕动症的程度 [15, 21, 29]。导致晕动症的其他因素还包括显示器、延迟、fl闪烁、卡顿、网络安全性等 [25, 38, 50, 93, 94]。晕动症检测领域的最先进研究大致可分为两类，分别采用主观方法、客观方法以及先进的机器学习/深度学习方法。此外，也有研究关注晕动症模型的降维。本节将依次讨论这些相关工作。

研究人员提出了多种主观测量方法来测量晕动症，例如模拟器晕动症问卷（SSQ） [8–10, 16, 84, 96], 快速晕动症量表（FMS） [43], 和晕动症易感性问卷（MSSQ） [42] 。与此相对，也有研究人员提出使用客观测量指标（即生理信号）来检测晕动症[31, 35, 73] 。先前研究表明，晕动症发生时，客观测量指标（例如心率、胃电快速心律失常、皮肤电反应、眨眼频率、瞳孔直径以及脑电图（EEG）的δ波和β波信号）会发生显著变化fi。例如，研究发现，心率和脑电δ波与晕动症呈正相关，而脑电β波则呈负相关 [14, 35, 56, 56, 77, 79] [56]。相反，另一项研究报告称，与其他可用于检测晕动症的客观测量指标相比，皮肤电反应与晕动症的正相关性更强 [35, 90]。

近年来，已有大量基于机器学习和深度学习的方法被提出 [1, 26, 31, 33, 35, 37, 44, 50, 68, 75, 78, 79, 95, 97,99] ，用于从多种主观测量（FMSQ、MSSQ）数据、客观测量（生理信号）数据以及集成多模态传感器测量（眼动追踪、头部追踪、运动流等）数据中，自动检测头戴式显示器中的晕动症。例如，Seungjun 等人 [1]提出了一种机器–深度–集成学习方法，用于根据生物生理数据对晕动症进行分类。相比之下， [73], 采用一种基于符号机器学习的方法来识别晕动症等级。此外，Azadeh 等人 [26] 采用基于支持向量机（SVM）的拓扑数据分析（TDA），并结合高斯径向基函数核方法，根据生理和主观测量数据预测晕动症。另一方面，Padmanaban 等人 [67]对虚拟现实视频数据中的人工设计特征应用机器学习算法，以预测晕动症。除了基于机器学习的方法，近年来，基于深度学习的方法也日益受到晕动症研究者的关注。例如， [41] 中的作者采用了三种基于深度学习/机器学习的方法，即卷积神经网络（CNN）、长短期记忆网络（LSTM）和支持向量回归（SVR），用于预测晕动症。相比之下，Lee 等人 [52] 使用三维卷积神经网络和多模态深度融合网络进行检测

利用虚拟现实视频数据中的光流、视差和显著性特征检测晕动症。另一方面，Jeong 等人[37] 采用基于注意力机制的深度学习模型，根据集成传感器数据预测晕动症。同样， [75], 采用基于长短期记忆网络的注意力网络，根据用户的生物生理信号检测晕动症。因此，Islam 等人 [35]采用基于长短期记忆网络的网络，根据用户的生物生理数据（例如，心率、皮肤电反应等）对晕动症严重程度进行分类。此外， [33] 提出了一种深度融合方法，根据多模态集成传感器数据（例如，眼动追踪、头部追踪等）对晕动症严重程度进行分类。尽管基于机器学习/深度学习的方法在晕动症检测方面已取得巨大成功，但将可解释人工智能用于解释检测出的晕动症模型仍存在显著的研究空白。事实上，理解为什么某些样本会被正确或错误地标记为晕动症，以及哪些特征促成了晕动症检测结果，是采取适当缓解措施的重要一步。因此，需要明确解释，以确定哪些特征（例如，眼动追踪、头部追踪、心率、皮肤电反应等）导致了晕动症检测。此外，在晕动症检测模型中应用可解释人工智能，可以显著提高模型的可信度，并深入揭示机器学习/深度学习模型得出特定fi决策的原因和方式。在本文语境中，可信指的是通过全局和局部解释、可解释层等机制，运用可解释人工智能解释晕动症，从而使机器学习/深度学习模型透明、易于理解，因而值得用户信赖 [92]。最近，Kundu 等人 [49] 采用了三种本质上可解释的机器学习模型，即可解释提升机（EBM）、决策树（DT）和逻辑回归（LR），根据用户的生物生理数据和主观测量数据检测并解释晕动症。然而，他们提出的方法仅限于二元分类（晕动症与无晕动症）..。相比之下，我们的工作针对晕动症检测考虑了多类别分类问题（无、低、中和高）。 [49] 中的另一项局限是，他们提出的 EBM 模型需要较长的训练时间 [66], ，因此不适用于实时部署。相比之下，我们的工作提出了一种轻量级超级学习器模型，并采用 SHAP、MSA、LIME 和 PDP 等事后解释方法，解释黑箱机器学习模型的晕动症检测结果。

目前已有一些关于晕动症模型降维的研究。许多研究人员采用基于 PCA 的方法识别重要特征并缩小模型大小 [45, 47, 60, 82, 85, 89]。例如，Lin 等人 [55]采用基于 PCA 的方法，从脑电信号中提取晕动症相关特征。提取特征后，他们使用了 3 种机器学习/深度学习模型，即线性回归（LR）、支持向量机（SVM）和自组织神经模糊推理网络（SOFIN），来预测用户的晕动症等级。同样， [48] 中的作者也使用 PCA 从脑电信号中 fi发现模式，并利用神经网络（NN）基于科罗拉多大学脑电数据集对认知任务进行分类。相较之下，Singla 等人在 [85] 使用 PCA 减少了模拟器晕动症问卷（SSQ）中的问题数量。另一方面，Mawalid 等人在 [61] 采用基于统计特征（如均值、变化量、标准差、峰值数量）和功率百分比频带的时域特征提取方法来了解晕动症特征，随后应用 K 近邻和朴素贝叶斯分类fi器对晕动症进行分类。然而，基于 PCA 的降维并非总是可靠，也存在一些缺点。例如，PCA 通过投影将高维数据映射到低维空间，这往往会导致原始数据[40]中的信息丢失。相比之下，我们采用基于可解释人工智能的方法对晕动症检测模型进行降维，以避免信息丢失并保持可信度。事实上，要开发有效的缓解方法，识别 VR 中诱发晕动症的关键特征至关重要。这可以通过使用可解释人工智能技术实现。然而，据我们所知，目前尚无研究探索用于检测和预测晕动症的可解释人工智能技术，这正是本文开展这项工作的动机。

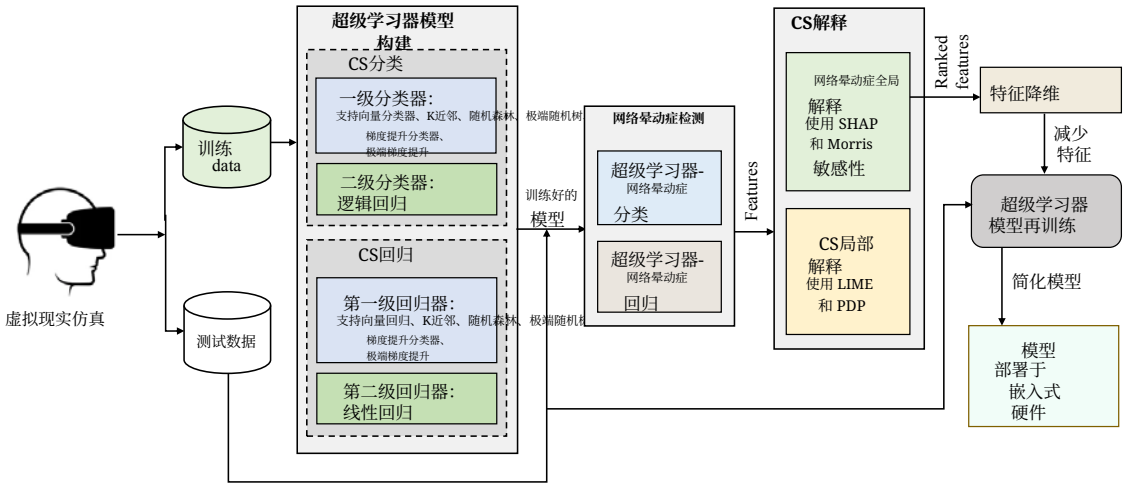


图 2. VR-LENS 用于晕动症（网络晕动症）检测、解释、减轻和部署的概览。

3 VR-LENS 框架

图 2 展示了所提出的 VR-LENS 框架概览，该框架用于 VR 晕动症检测、解释、模型缩减和部署。首先，使用训练数据训练基于超级学习器的集成学习模型，以完成分类和回归任务的训练（详见下一节）。晕动症分类和回归训练阶段包含两个分类器层级。第 1 层分类器由多个机器学习模型作为基模型组成，第 2 层分类器则由一个元分类器组成，用于完成 fi最终分类。接下来，使用训练好的基于 VR-LENS 的分类和回归模型，对测试数据集中的晕动症进行分类和回归。晕动症回归旨在使用训练好的超级学习器模型对晕动症等级进行分类。相比之下，晕动症回归预测当前晕动症 FMS 评分的下一个值，其范围为 0 到 10。下一阶段使用 SHAP、LIME、莫里斯敏感性分析和 PDP 等事后解释方法，解释晕动症预测结果。对于全局解释，我们使用 SHAP 和莫里斯敏感性工具，根据整体结果解释整体特征排序。相反，局部解释则基于 LIME 和 PDP，针对特定的fi测试样本对各个特征进行排序。通过解释方法分析并排序特征后，在模型缩减阶段，我们仅使用排名靠前的特征（在本例中为特征的 1/3）重新训练超级学习器模型。这样可以显著缩小模型，并减少超参数数量，从而缩短训练时间和推理时间。最后，我们将精简超级学习器模型部署到嵌入式平台上，用于晕动症检测。

3.1 使用超级学习器进行晕动症分类

我们实现了一种基于超级学习器的集成模型，用于组合基模型的预测结果，并利用外生变量的信息增强预测效果。值得一提的是，超级学习方法基于多个集成算法的通用框架 [101]。我们首先构建基学习器，并使用元学习器对基学习器进行 fi。我们使用六个堆叠的基学习器，并分别进行测试，以获得最佳性能。其思路是选择一个次优分类fi器来解决问题，提升预测性能，并提高

算法1基于超学习的晕动症分类

输入： 训练数据集： X_{Train},
测试数据集： X_{Test} X_{Test},
基学习器列表B_L,
元学习器M_L;
输出： D_L ={ }
1：**对于每个**基学习器 $i \in B_L$ **执行**
2: $F_L = \lambda(X_train)$
3: $X_train' = \theta(X_train, F_L)$
4: **结束循环**
5： $F_L' = \lambda(\mathrm{M}_L, X_train')$
6： $P_L = \theta'(F_L', X_test)$
7： $D_L = D_L - \mathrm{P}_L$
8: **返回** D_L

超级学习器模型的泛化性能。我们用于构建超级学习器模型的基学习器包括：支持向量分fi器（SVC） [12], K近邻分类fi器（KNNC）[4],极端随机树分类fi器（ETC）[23],XGBoost分类fi器（XGBC）[11],随机森林分类fi器（RFC）[54], 梯度提升分类fi器（GBC）[18], ； 元学习器分fi器则基于逻辑回归分fi器 [13]。

算法1展示了用于晕动症检测的基于超级学习器的集成机器学习方法的概览。该算法以若干基学习器、元学习器以及训练和测试数据集为输入，并返回晕动症等级。首先，对于每个基学习器 B_L，使用学习算法函数 λ 以堆叠方式对训练数据 X_train 进行拟合，从而获得最佳拟合的基学习器，记为 F_L （第2行）。接下来，基于预测函数 θ 创建元学习训练集X_train' （第3行）。这将使用 X_train和 F_L为元学习器生成新的数据集。随后，在训练函数 λ 中对元学习器 M_L和元学习训练集 X_train'进行拟合（第5行）。 λ 函数根据集成基学习器和元学习器计算训练后的值 F_L。然后，基于预测算法 θ，使用拟合后的元学习器 F_L' 预测晕动症等级 D_L （第6–7行）。最后，在对所有测试数据完成此过程后，算法返回晕动症等级（第8行）。

3.2 使用超级学习器进行晕动症回归

我们在晕动症回归任务中使用了与分类任务相同的超级学习器模型，其基学习器包括支持向量回归器（SVR）[12], K近邻回归器（KNNR）[4], 极端随机树回归器（ETR）[23], XGBoost回归器（XGBR）[11], 随机森林回归器（RFR）[54], 梯度提升回归器（GBR）[18]。随后，我们使用元学习器回归器，即线性回归模型，预测范围为0到10的晕动症FMS评分。通过比较用户当前的生理状态与先前的生理状态，晕动症回归可对用户正在经历的晕动症FMS评分进行回归预测。晕动症回归任务可形式化定义如下：给定观测到的VR数据历史（例如，眼动追踪数据、头部追踪数据、生理信号、游戏玩法数据等）以及先前时间步的FMS评分 t 1，预测下一个时间步的FMS评分 t。基于先前时间 t 1的生理信号，对时间 t 的晕动症进行的预测记为 CSR_t。例如，如果我们要预测时间为 t = 20 秒时的晕动症FMS评分，那么 CS_t 可表示为：

$$CSR_t \Rightarrow [\ P_{t-19}, P_{t-18}, P_{t-17}, P_{t-16} \cdots, P_t]$$

表1。 非简化超级学习器模式下，10折交叉验证在晕动症严重程度分类中的性能 l)

数据集	% 准确率	精确率				召回率				%F1 分数			
		无 低		中	高 无 低			中	高 无 低			中	High
游戏玩法	82	91	70	56	75	92	90	67	65	91	79	61	71
生物生理	98	95	-	96	95	96	-	94	97	96	-	96	95

表2. 集成传感器数据集上晕动症严重程度分类的10折交叉验证性能（未简化的超级学习器模型）

模态融合	% 准确率	% 精确率				% 召回率				%F1 分数			
		无 低		中	高 无 低			中	高 无 低			中	High
头部追踪	67	71	63	43	64	79	71	48	40	75	67	42	49
眼动追踪	93	97	91	77	90	96	96	68	81	97	94	73	86
眼动 +头部追踪	95	98	95	87	94	99	97	79	90	98	96	83	92

其中 P 表示用户的认知状态。

3.3 晕动症解释

用于解释超级学习器模型所产生的晕动症检测结果的可解释人工智能工具，会生成可视化表示，形式可以是条形图，也可以是一组全局或局部解释的可视化图。我们利用这些图表来了解晕动症检测与预测中的模型可解释性。解释可分为全局和局部解释。晕动症检测的整体重要性排序（全局解释）以条形图形式呈现。对于局部解释，我们从测试数据集中随机选取每个样本，其中包含所有特征。我们使用 SHAPley加性解释（SHAP）和Morris敏感性分析（MSA）来解释整体特征重要性。SHAP 根据博弈论方法确定特征重要性。另一方面，莫里斯敏感性分析通过一次调整一个特征来衡量其影响，并根据这一随机化过程确定特征重要性。我们使用全局解释的结果来识别关键特征，从而有效缩减特征空间。

对于局部解释，我们使用局部可解释模型无关解释（LIME）和部分依赖图（PDP）工具，解释测试样本的个体预测。LIME 通过用可解释模型近似底层模型来生成解释，从而展示单个样本中的哪些特征促成了输出。同样，PDP 展示一个或两个特征对机器学习模型预测结果的边际 efffect。该边际 efffect 可能导致输出与特征之间呈线性、单调或更复杂的关系。

4 数据集与实验设置

本节介绍我们的实验设置和数据集，以验证所提出的 VR-LENSE框架。我们使用 Scikit-Learn [71] 训练并评估所提出的基于超级学习器的机器学习模型。为解释该超级学习器模型，我们使用了 SHAP [57] 和 InterpretML [66] 库。为了将基于超级学习器的机器学习模型部署到搭载 Qualcomm Snapdragon 750G处理器的 Android设备上，我们使用了 Android Studio [17], 和 ONNX [5]库。该基于超级学习器的机器学习模型在配备英特尔酷睿i9处理器、32GB 随机存取存储器和 NVIDIA GeForce RTX 3080 Ti GPU 的设备上进行训练。

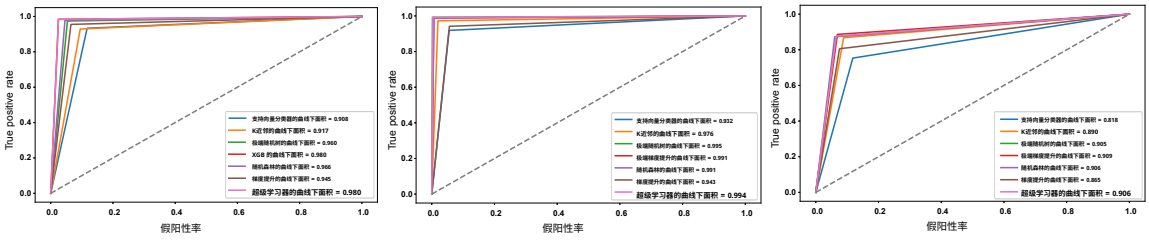
4.1 数据集

为验证所提出的 VR-LENS 框架在晕动症分类、回归、解释和部署方面的有效性，我们使用了三个数据集，包括集成传感器生物生理数据集和游戏玩法数据集。

4.1.1 集成传感器数据集。集成传感器数据集 [33]包含 27 名参与者（男性：15 名，女性：12 名）沉浸于 5 种不同 VR 模拟中的眼动追踪、头部追踪和生理信号数据：Beach City、Road Side、过山车、SeaVoyage 和家具店。研究共招募了 30 名参与者（男性：15 名，女性：15 名）来收集实验数据。然而，由于技术问题（即黑牙和 HTC-Vive 无线适配器黑屏问题），有三名参与者的数据未能采集。眼动追踪、头部追踪和生理数据包含不同的子类别。例如，眼动追踪数据的子类别包括瞳孔直径（左）、瞳孔位置（x、y、z）、注视方向（x、y、z）、辐辏距离和眼睛睁开程度百分比；头部追踪数据的子类别则分别为 X、Y、Z 和 W 轴的四元数旋转。类似地，生理信号的子类别包括皮肤电活动（EDA）和心率（HR）测量值。该数据集共包含 20104 个样本，记录时长最长为 7 分钟的 VR 仿真。此外，数据集包含四种不同的晕动症严重程度类别：无、低、中和高；FMS评分范围为 0 至 10，用于回归分析。

4.1.2 生物生理数据集。生物生理[36] 数据集包含多种生理信号，例如心率（HR）、呼吸频率（BR）、心率变异性（HRV）和皮肤电反应（GSR）。HR、BR、GSR 和 HRV 数据各有不同的子类别。例如，HR 数据的子类别包括相对于静息基线的变化百分比（PC）、3s 滚动窗口内的最小值（MIN）、3s 滚动窗口内的最大值（MAX）以及 3s 滚动窗口的移动平均值（AVG）。其他生物生理信号的子类别也相同。研究人员从一所大学的班级中招募了共 31 名大学生参加实验（男性：29 名，女性：2 名）。遗憾的是，由于一些原因（例如，实验过程中 HR 传感器的电池耗尽、部分用户在实验过程中感到严重晕动症等），他们未能采集 8 名用户的数据。因此，这些生理信号采集自沉浸于虚拟过山车模拟中的 23 名参与者。该数据集共包含 24533 个样本，记录的虚拟现实仿真时长最长为 897 秒。该数据集包含三种不同的晕动症严重程度类别：轻度晕动症、中度晕动症和急性晕动症。为减少歧义，我们将“轻度晕动症”标记为“无”，“中度晕动症”标记为“中”，并将“急性晕动症”标记为“高”晕动症类别。用于回归分析的 FMS 评分范围为 0 到 10。

4.1.3 游戏数据集。游戏数据集 [74] 包含来自多个来源的 22 种不同特征，例如候选对象档案、问卷、用户视野、用户位置、游戏进行时的游戏速度等，涵盖 87 名参与者。该数据集使用两款虚拟现实游戏生成，即赛车游戏和飞行游戏。然而，由于只有 35 名参与者（男性：26 名，女性：9 名）的晕动症数据有效，因此仅收集了这些参与者的数据。数据来自游戏中正确回答所有虚拟现实主观问卷（VRSQ）并完成全部游戏交互的参与者。该数据集包含四种晕动症严重程度类别：无、轻微、中度和严重。FMS评分范围为 0 到 10，用于回归分析。该数据集共包含 9391 个样本，记录于 5 分钟的 VR 游戏模拟过程中 [73]。值得一提的是，为了减少歧义，我们将这四个类别重新命名如下：无：无，轻微：低，中度：中，严重：高，与集成传感器数据集的命名方式一致。



(a) (b) (c)

图3. 超级学习器的受试者工作特征曲线（**a**）集成传感器数据集（**b**）游戏玩法数据集。（**c**）生理数据集。

4.2 部署设置

我们使用一台运行 Android 12 的 Samsung A52 5G 设备评估所提出的超级学习器模型。该设备搭载高通骁龙750G处理器，最高时钟频率为 2.2 GHz，配备 6 GB 随机存取存储器和 128 GB 存储空间。高通骁龙处理器常用于虚拟现实头戴式显示器，因此我们选择了这款设备。在此设置中，我们通过注入集成传感器数据来模拟虚拟现实环境，使模型检测晕动症严重程度，并测量模型的输出和推理时间，以评估其性能。

4.3 性能指标

采用超级学习器的集成模型对晕动症类别进行分类，其性能使用准确率、精确率、召回率、F1分数、曲线下面积（AUC）和受试者工作特征曲线（ROC） [1]等标准质量指标进行评估。同样，回归模型的性能使用均方根误差（RMSE）、平均绝对误差（MAE）、皮尔逊线性相关系数（PLCC） [102], 和 R^2 分数等常用损失函数进行分析。例如，若 y_t 和$\hat{}y_t$ 分别表示候选对象在时间 t的实际和预测晕动症，则RMSE可定义如下。

$$RMSE = \sqrt{\frac{1}{|N|}\sum_{y\in S}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}$$

其中， N 和 S 分别表示样本总数和时间步数。从数学上讲，MAE 可表示如下。

$$MAE = \frac{1}{|N|}\sum_{y\in S}\sum_{t=1}^{N}(y_t - \hat{y}_t)$$

值得一提的是，MAE 和 RMSE 越小，回归模型的表现越好。若 Õ$(y_t \ \hat{}y_t)^2$ 和 Õ$(y_t \ \bar{y})^2$ 分别表示回归平方和（SSR）和总平方和（SST），那么 R^2 分数可表示如下。

$$R^2 = 1 - \frac{\sum(y_t - \hat{y}_t)^2}{\sum(y_t - \bar{y})^2}$$

因此 目前，低 R^2 值表明回归模型未能充分捕捉输出方差 e.

数据集	MAE	R^2	RMSE	PLCC
游戏玩法	0.11	0.57	0.21	0.75
生物生理	0.01	0.99	0.02	0.99

（a）对于生物生理和游戏玩法数据集

模态融合	MAE	R^2	RMSE	PLCC
头部追踪	0.72	0.49	1.08	0.70
眼动追踪	0.23	0.92	0.43	0.96
眼动 + 头部追踪	0.02	0.92	0.04	0.96

(b) 集成传感器数据集

表3. 使用未简化的超级学习器模型进行晕动症回归（包含所有特征）

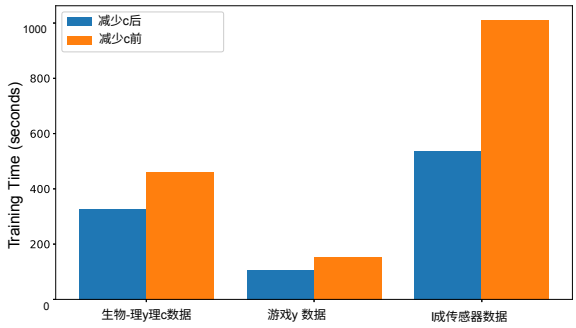


图 4. 超级学习器模型的训练时间：未简化模型（全部特征）与简化模型（减少后的特征）对比。

5 结果

本节介绍晕动症检测、解释、减轻和部署的结果 t.

5.1 使用全部特征时的晕动症分类与回归性能

在应用基于可解释人工智能的模型缩减之前，我们先介绍未简化的超级学习器模型开发结果及其重要统计数据，例如训练时间和推理时间、晕动症分类以及回归准确率。这有助于我们在论文接下来的章节中比较所提方法（基于可解释人工智能的缩减）的有效性。

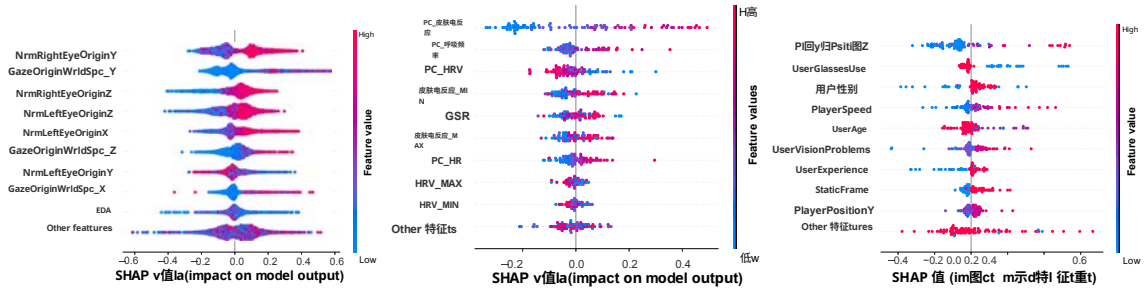
表4 超级学习器模型的超参数（未减少版与减少版）

模型		超参数
KNN	原始	默认
	减少	邻居数量 = 2
SVC	原始	默认
	减少	核= 多项式
ET	原始	默认
	减少	估计器数量 = 50
RF	原始	默认
	减少	估计器数量 = 10
XGB	原始	默认
	减少	估计器数量=30
GB	原始	默认
	减少	学习率= 0.05

5.1.1 晕动症检测模型开发。表4列出了基于超级学习器的晕动症检测模型的超参数。模型使用 Scikit-learn 的默认超参数进行训练 [72] ，其中每个学习器均采用默认超参数值。我们采用 10 折交叉验证技术训练超级学习器模型，并测试其性能，方式与 [6] 类似：将数据集划分为 k 组（即本研究中为 $k = 10$）。其中仅使用一个分区测试模型，其余分区用于训练。该方法重复 k 次，每次选取一个新的测试分区，并将其余（k 1）个分区作为训练数据集，以消除偏差。请注意，我们没有采用留一参与者验证技术来训练所提出的超级学习器模型并测试其性能。这是因为我们的数据集相当大，而留一参与者验证更适用于小型数据集，因为它在每次迭代中会使用更多训练样本 [58]。集成传感器、游戏玩法和生理数据集的训练时间与推理时间如图4所示。使用超级学习器（包含所有特征）训练集成传感器、游戏玩法和生理数据集分别需要 1012 秒、460 秒和 154 秒。

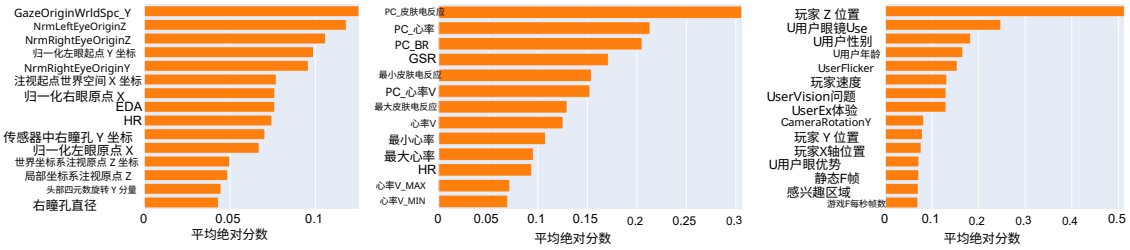
5.1.2 晕动症分类。我们在表1和表2中总结了使用超级学习器模型进行晕动症严重程度类别分fi类的准确率、精确率、召回率和F1分数。对于集成传感器、生物生理和游戏玩法数据集，我们的超级学习器模型分别达到了95%、98%和82%的准确率。从表1可以观察到，就晕动症严重程度分fi类的精确率、召回率和F1分数而言，超级学习器模型在生理数据集上的整体性能显著优于游戏数据集。例如，生理数据集在无、 中和高晕动症类别上的精确率分别为95%、96%和95%。另一方面，游戏数据集在无、低、中和高晕动症类别上的精确率分别为91%、70%、56%和75%。如第4.1.2节所述，生物生理数据集仅包含3个晕动症类别，即无、中和高。因此，低晕动症类别没有相应的结果。此外，图3展示了使用超级学习器模型对集成传感器、生理和游戏数据集得到的AUC-ROC曲线。从图3b和图3c可以观察到，生物生理数据集的AUC分数较高，为0.994；然而，游戏数据集的AUC分数相对较低，为0.906。此外，与其他基线模型相比，所提出的超级学习器模型在所有数据集上均取得了更高的AUC分数。例如，对于生物生理数据集，超级学习器的AUC分数为0.994，而SVC的AUC分数为0.932。可以观察到，除生物生理和游戏数据集中的XGB分类fi器模型外，超级学习器的性能优于其他基线模型。这是因为集成方法通过使用元学习算法 [81]组合一组独立的学习算法，来近似数据中的复杂函数关系。这能够深入挖掘特征信息，从而提高分fi类准确率。同样，对于生物生理数据集，我们所提出的超级学习器模型的准确率也优于基于深度学习的卷积LSTM模型 [31]。尽管我们在游戏数据集上取得了较高的准确率，但无法将其与基于符号机器学习的方法 [73]进行比较。原因在于，这些方法执行的是二分类，而我们对晕动症严重程度进行的是分类（多分类fi）。

为展示所提出模型的性能，并将我们的工作与基于最先进深度学习模型的晕动症检测进行比较，在 [33], 中，我们使用集成传感器数据集中的三种模态融合方式（头部追踪、眼动追踪和头部 + 眼动追踪），如表2所示。所提出的超级学习器模型在头部追踪、眼动追踪以及头部 +眼动追踪 模态融合方式上的准确率分别达到 67%、93%和95% 。在眼动追踪 和头部 +眼动追踪 模态融合方式上，该模型的表现优于 [33] 中提出的先前深度融合模型，其准确率分别为80.7% 和87.7%。同样，图3a展示了以下数据的 AUC-ROC 曲线



(a) (b) (c)

图 5. 使用 SHAP 全局解释的总体特征重要性：（**a**）集成传感器数据集；（**b**）生物生理数据集；（**c**）游戏数据集。



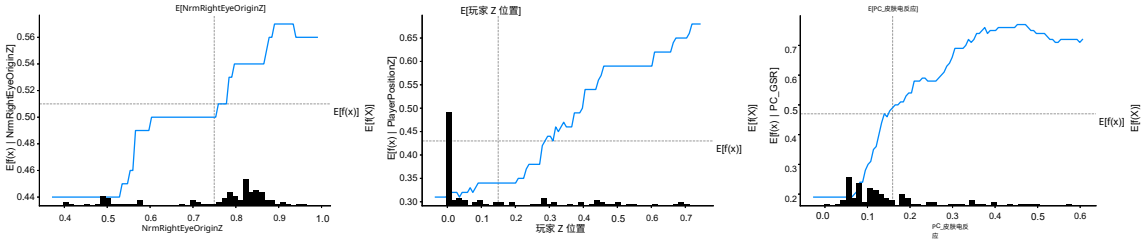
(a) (b) (c)

图 6. 用于基于 MSA 的全局解释的总体特征重要性：**(a)** 集成传感器数据集 **(b)** 生物生理数据集 **(c)** 游戏数据集。

集成传感器数据集采用头部 + 眼动追踪 模态融合方式，这是因为其准确率较高。可以观察到，集成传感器的 AUC 分数为 0.980，高于其他单一分类器，fi但 XGB 分类器模型除外。例如，超级学习器的 AUC 分数较高，为 0.980；然而，基线模型 SVC 的 AUC 分数相对较低，为 0.908。

5.1.3 晕动症回归。表 3a 和表 3b 总结了使用超级学习器模型对集成传感器、生物生理和游戏数据集进行晕动症回归时的平均绝对误差、决定系数、均方根误差和皮尔逊线性相关系数。生物生理和游戏数据集的平均绝对误差、决定系数、均方根误差和皮尔逊线性相关系数分别为 0.01、0.99、0.02、0.99 和 0.11、0.57、021、0.75。与分类类似，回归也采用三种模态融合方式（头部追踪、眼动追踪和头部 + 眼动追踪），如表 3b 中集成传感器数据集所示。

所提出的超级学习器模型在晕动症回归方面优于此前报告的结果 [33] 。例如，深度融合模型的回归结果显示，其 R^2 得分值分别为0.18、0.56和0.67；而所提出的超级学习器模型在头部追踪、眼动追踪以及头部 + 眼动追踪模态融合方面的 R^2 得分值分别为0.49、0.92和0.92 。值得一提的是，高 R^2 值表明该回归模型能够很好地对持续发生的晕动症进行回归。超级学习器模型表现良好的原因在于，它基于集成技术构建，从而显著fi提升了模型性能。



(a) (b) (c)

图 7. 基于 PDP 的晕动症分类局部解释：（**a**）集成传感器数据集。（**b**）游戏数据集。（**c**）生理数据集。

5.2 基于可解释人工智能的晕动症解释模型缩减

在本节中，我们应用事后解释方法，即 SHAP、莫里斯敏感性分析、LIME 和 PDP，来解释超级学习器模型结果。随后，我们利用这些解释，特别是 SHAP 和 MSA 的全局解释，识别主导特征，以改进基于超级学习器的晕动症检测模型。

5.2.1 晕动症全局解释。图5基于SHAP解释，展示了使用超级学习器模型，针对生理、游戏玩法和集成传感器数据集中的全部特征进行晕动症严重程度分类时的总体特征重要性。SHAP值用于计算对晕动症严重程度分类贡献最大的特征排名，最重要的特征位于顶部，最不重要的位于底部。从图5a可以看出，对于集成传感器数据集，NrmRightEyeOriginY（对应Y轴测量中的归一化右眼起点）、GazeOriginWrldSpc_Y（对应世界空间中Y轴上的注视起点）和NrmRightEyeOriginZ（对应Z轴上的归一化右眼起点）等特征，是晕动症严重程度分类中最主要的特征。值得一提的是，相较于头部追踪特征，眼动追踪特征对晕动症严重程度分类的影响要强得多。这是因为眼动追踪特征包含丰富的信息，例如用户眨眼的类型、注视行为以及用于追踪用户活动的瞳孔位置 [33, 34, 37]。同样，从图5b可以看出，对于生理数据集，PC_GSR（对应皮肤电反应（GSR）测量值的百分比）、PC_BR（对应呼吸频率（BR）测量值的百分比变化）和PC_HRV（对应心率变异性（HRV）测量值的百分比变化）等特征，对晕动症分类的影响要强得多。同样，对于游戏数据集，Player Position Z（玩家 Z 位置）、用户眼镜佩戴情况和用户性别等特征，是晕动症分类中最重要的特征（图5c）。图6展示了使用Morris敏感性分析对集成传感器、生理和游戏数据集进行晕动症分类时的总体特征重要性。平均绝对分数（MAS）用于计算对晕动症分类贡献最大的特征排名。从图6a可以看出，对于集成传感器数据集，对晕动症分类有贡献的大多数特征仍是眼动追踪特征，即GazeOriginWrldSpc_Y（对应世界空间中Y轴测量上的注视起点）、NrmLeftEyeOriginZ（对应Z轴测量上的归一化左眼起点）、NrmRightEyeOriginZ（对应Z轴测量上的归一化右眼起点）等。同样，从图6b和6c可以看出，最能预测晕动症严重程度分类的特征

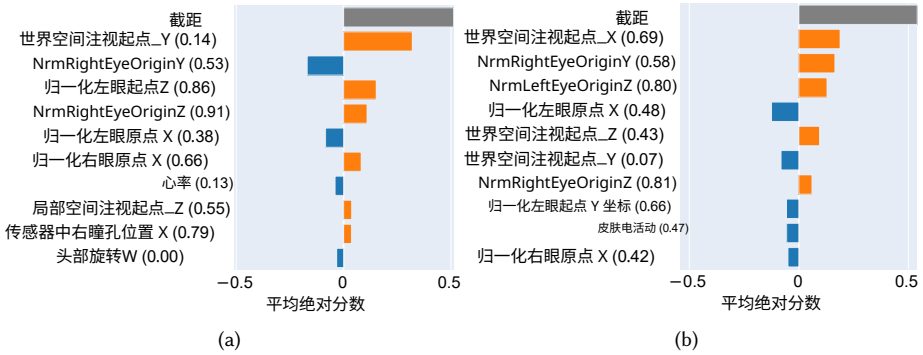


图8. 基于 LIME 的整合数据集晕动症分类局部解释：(a) 晕动症严重程度高时的解释；(b) 晕动症严重程度低时的解释。

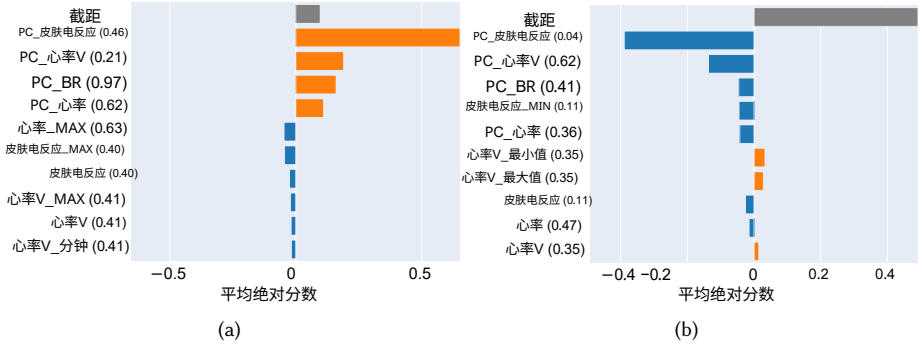


图 9. 基于 LIME 的生理数据集晕动症分类局部解释：(a) 高晕动症严重程度的解释；(b) 低晕动症严重程度的解释。

对于生理数据集而言，是*PC_GSR*、PC_HR、*PC_BR*等；而对于游戏数据集而言，则是玩家 *Z* 位置、用户眼镜佩戴情况、用户性别等。

5.2.2 晕动症局部解释。利用 LIME 对集成传感器数据集进行局部解释的结果如图 8 所示。图 8a 展示了高晕动症严重程度类别的分类，其中黄色和黑色的条形表示该个体结果对应的晕动症概率 MAS 值。 *x*轴表示模型输出的 MAS 值，即对数优势（特征重要性在预测中的概率）； *y*轴列出了模型的特征。大多数特征都产生负向影响，以黄色条形表示；因此，模型能够准确判定高晕动症严重程度类别。观察发现，眼动追踪特征世界空间注视起点_*Y*（对应世界空间中注视原点的 Y 轴测量）是分类高晕动症严重程度类别时影响最大的特征，其 MAS 值最高，接近 0.4。例如，在图 8b 中，无晕动症严重程度类别的分类对大多数特征呈现负 MAS 值，这表明大多数特征产生正向影响。除皮肤电活动（对应皮肤电活动测量）外，大多数特征都属于眼动追踪特征；因此，模型能够恰当地判定无晕动症严重程度类别。类似地，生理数据集的晕动症分类局部解释如图9所示

在图9中。在图9a和图9b中可以看到，大多数特征都支持高晕动症严重程度类别，例如*PC_*皮肤电反应 （对应皮肤电反应测量值的百分比）、*PC_*心率变异性 （对应心率变异性测量值的百分比）等；而对于无晕动症严重程度类别，主导特征包括心率变异性_最小值 （对应最小心率变异性测量值）、心率变异性_最大值 （对应最大心率变异性测量值）等。同样，游戏数据集的晕动症分类局部解释如图10所示。从图10a可以看出，对高晕动症严重程度类别影响最大的特征属于眼动追踪特征，例如用户性别、用户年龄等。因此，图10b显示，对正向结果（无晕动症严重程度类别）产生影响的大多数特征包括玩家 Z 位置、用户年龄、静态帧、用户视力问题等。如第5.1.2节所述，游戏数据集的晕动症分类准确率较低。因此，诸如用户年龄和玩家位置 Y 等特征在正向和负向结果中都出现了错误解释。这类特征的局部解释有助于深入理解分类/误分类结果，从而增强对模型结果的信任，以便作出恰当决策。

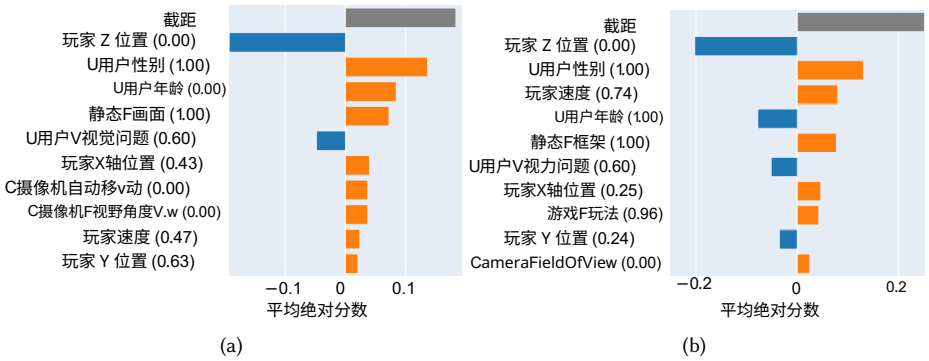


图 10. 基于 LIME 的游戏数据集晕动症分类局部解释：(a) 高晕动症严重程度的解释；(b) 低晕动症严重程度的解释。

我们利用基于 PDP 的局部解释分析重要特征与模型输出之间的关系，以深入了解晕动症严重程度类别分类的解释。图 7 展示了集成传感器、生理数据集和游戏数据集的重要特征的 PDP 结果。PDP 解释旨在确定给定的一组特征与相应预测值之间的部分关系。图 7a 显示，眼动追踪特征 *NrmRightEyeOriginZ*（对应于 Z轴 测量中的归一化右眼原点）与晕动症分类呈正相关（该值越高，发生晕动症的概率越高）。同样，图 7b 和 c 显示，游戏数据集中的特征*z*轴上的玩家位置 以及生理数据集中的 *PC_*皮肤电反应特征与晕动症分类呈正相关。

本节　　介绍特征选择、晕动症模型缩减，以及在嵌入式平台上的部署。

5.2.3 面向部署的模型缩减。利用上一节所述的基于 MSA 和 SHAP 的（全局）解释结果，我们首先确定排名前 1/3 的特征，并重新训练基于超级学习器的集成模型。对于我们的超级学习器模型和数据集，1/3 的比例是通过反复试验法获得的。例如，我们从集成传感器数据集中选取排名前 15 个特征（共 43 个），从生理数据集中选取排名前4 个特征（共 13 个），并从游戏过程数据集中选取排名前 10 个特征（共 20 个）。表6展示了

表 5。 超级学习器部署的推理时间：减少前（所有特征）与减少后（减少后的特征）)

样本量	推理时间（秒）	
	特征降维前	特征降维后
1	0.15	0.07
10	1.2	0.72
100	11.8	7.6
500	54.7	38.04

表6. 选定特征列表 用于通过全局解释重新训练超级学习器模型的 全局解释

数据集	基于 SHAP 和 MSA 的全局解释所选特征
集成传感器	NrmRightEyeOriginZ；NrmLeftEyeOriginY；NrmRightEyeOriginX；NrmRightEyeOriginY；NrmLeftEyeOriginZ；NrmLeftEyeOriginX；NrmSRLeftEyeGazeDirX；NrmSRLeftEyeGazeDirY；NrmSRRightEyeGazeDirY；世界空间注视起点_Y；世界空间注视起点_Z；右瞳孔直径；头部四元数旋转 W 分量；头部四元数旋转 Y 分量；头部欧拉角 X 分量
游戏玩法	玩家 Z 位置；用户眼镜使用情况；用户性别；用户年龄；玩家速度；用户视力问题；用户经验；玩家X轴位置；静态帧；摄像机旋转 Y 分量
生物生理	PC_皮肤电反应；PC_呼吸频率；皮肤电反应_最小值；PC_心率变异性

表7. 使用精简超级学习器模型（减少特征）进行晕动症严重程度分类的10折交叉验证性能

数据集	特征数量		准确率	精确率				召回率				F1分数			
	原始	减少		无	低		中	高	无	低		中	高	无	低		中	High
游戏玩法	22	10	81	88	81	79	94	95	89	75	80	91	79	78	80
集成传感器	43	15	96	99	95	91	95	98	98	97	95	99	96	89	94
生物生理	14	4	98	99	-	97	99	99	-	98	96	99	-	98	99

通过 SHAP 和基于 MSA 的全局解释选出的特征，这些特征用于重新训练所提出的基于超级学习器的集成模型。同样，表4列出了重新训练的超级学习器模型的超参数。

5.2.4 精简超级学习器模型的性能。 部署在 Samsung A52 5G 设备上的模型总大小，未简化超级学习器模型为 133,609 KB，精简超级学习器模型为 90,917 KB。图4展示了在精简后的生物生理、游戏玩法和集成传感器数据集上使用精简超级学习器模型所带来的训练时间改善。可以看到，对于生物生理、游戏玩法和集成传感器数据集，精简超级学习器模型的训练时间分别显著提升了 1.53 倍、1.46 倍和 1.91 倍。表5列出了部署后的精简和未简化超级学习器模型的推理时间。例如，在部署后的未简化超级学习器模型上，对 1、10、100 和 500 个样本进行推理分别需要 0.15、1.2、11.8 和 54.7 秒。相比之下，精简超级学习器模型处理 1、10、100 和 500 个样本仅需 0.07、0.72、7.6 和 38.04 秒，速度分别比未简化超级学习器模型快 2.15 倍、1.67 倍、1.58 倍和 1.44 倍。

表 8. 使用精简超级学习器模型（采用减少后的特征）对生物生理、集成传感器和游戏数据集进行晕动症回归

数据集	MAE	R^2	均方根误差	皮尔逊线性相关系数
生物生理	0.03	0.95	0.06	0.97
集成传感器	0.02	0.95	0.03	0.97
游戏玩法	0.12	0.49	0.22	0.70

5.2.5 精简超级学习器模型的晕动症分类。 如前所述，我们将超级学习器模型部署在嵌入式硬件上，用于对集成传感器数据集进行分类。随后，我们在其余数据集上模拟超级学习器模型，以评估其性能。

表7总结了使用降阶超级学习器模型对集成传感器、生物生理和游戏玩法数据集进行晕动症类别分类的准确率、精确率、召回率和F1分数。例如，使用精简超级学习器模型对集成传感器数据集进行晕动症类别分类，准确率达到96%。此外，生理数据集的晕动症类别分类准确率为98%，略高于该数据集使用未简化的超级学习器模型时的准确率（见表1）。然而，与未简化版本相比，游戏数据集使用精简超级学习器模型时的晕动症类别分类准确率略微下降了1.2%。此外，对于这三个数据集，与未简化版本相比，精简超级学习器模型在无、低、中和高晕动症类别上的其他性能指标（如精确率、召回率和F1分数）略有提高。例如，使用集成传感器数据集的精简超级学习器模型，在无、低、中和高晕动症类别上的精确率分别为99%、95%、91%和95%，略优于未简化的超级学习器模型。同样，精简超级学习器模型在无、中和高晕动症类别上的召回率分别为99%、98%、97%和95%，其中中等类别比未简化模型高1.23。此外，对于其他数据集，特别是生理数据集，简化模型提高了所有类别的精确率、召回率和F1分数。

5.2.6 精简超级学习器模型的晕动症回归。表 8 展示了使用降阶超级学习器模型对集成传感器数据集、生理数据集和游戏数据集进行晕动症回归的性能。例如，集成传感器数据集的精简超级学习器模型对应的平均绝对误差、(R^2)、均方根误差和皮尔逊线性相关系数分别为 0.02、0.95、0.04 和 0.97；生理数据集对应的值分别为 0.03、0.95、0.06 和 0.97。该用于集成传感器数据集和生理数据集的精简超级学习器模型显著提高了平均绝对误差、(R^2)、均方根误差和皮尔逊线性相关系数。另一方面，对于游戏数据集，平均绝对误差、(R^2)、均方根误差和皮尔逊线性相关系数分别为 0.12、0.49、0.22 和 0.70，略优于未简化模型。这是因为游戏玩法数据包含的特征大多是类别特征。.

6 讨论

本节简要讨论使用 VR-LENS 框架获得的结果。SHAP 和基于 MSA 的全局解释揭示，对于集成传感器数据集，正常眼睛原点、注视原点、瞳孔直径等特征，是导致晕动症最具影响力的特征。同样，对于生物生理和游戏玩法数据集，*PC_*皮肤电反应、*PC_*呼吸频率 和 *PC_*心率变异性等特征，以及玩家 Z 位置、用户眼镜佩戴情况、用户性别等特征，分别发挥了重要作用。相反，LIME 和基于 PDP 的局部解释为具体预测提供了有用见解。因此，我们根据 SHAP 和 MSA 的全局解释对特征重要性进行了排序，

并使用重要特征重新训练超级学习器模型。我们的结果表明，由 SHAP 和 MSA 指导的精简超级学习器模型显著缩短了训练时间。此外，在基于高通骁龙750G处理器的 Samsung A52 5G 设备上部署的精简超级学习器模型，在不牺牲准确率的情况下，可更快地推理，实现实时晕动症检测。例如，部署的精简超级学习器晕动症模型仅使用非精简版本三分之一的特征进行训练，便以 96% 的准确率对晕动症严重程度进行了分类。同样，在对晕动症进行回归时，精简超级学习器在集成传感器数据集上取得了 0.03 的均方根误差，比非精简版本低 25.5%。值得一提的是，精简超级学习器在生理数据集上的晕动症分类和回归任务中均表现良好。相比之下，游戏数据集在晕动症分类和回归任务中的表现都较差。原因在于，游戏数据集所含特征大多来自用户档案数据，无法提供关于晕动症的有用见解。然而，眨眼频率、瞳孔直径、心率等集成传感器数据，为了解用户与晕动症相关的行为提供了更多信息。

我们提出的超级学习器模型的准确率优于若干基于机器学习和深度学习的晕动症检测最先进研究。例如，Islam 等人 [34] 使用深度时序卷积网络（DeepTCN），根据眼动追踪、心率和皮肤电反应数据预测晕动症快速晕动症量表（FMS）评分（范围为0–10），其均方根误差为0.49。相比之下，Dennison等人 [14] 报告称，基于生物生理数据，准确率为78%， R^2 值为75%。我们超级学习器模型的分类和回归准确率优于这些研究。此外，还有其他与本研究相关的工作。例如，Qu等人 [75], 、Kim等人 [46], 、Garcia-Agundez等人 [20], 和 Jeong等人 [39] 分别利用生物生理数据和脑电图/心电图信号报告了96.85%、、89.16%、82%和94.02%的晕动症检测准确率。

尽管已有多项关于晕动症检测方法的研究，但迄今只有少数研究探讨了晕动症的成因 [31, 33, 46, 67]。然而，据我们所知，尚无研究将可解释人工智能（XAI）应用于解释黑箱机器学习模型检测晕动症的结果、利用 XAI 缩减机器学习模型大小，并将其部署到嵌入式设备上。事实上，基于 XAI 的解释有助于研究人员理解晕动症分类正确或错误的原因 fi，还可进一步用于开发有效的晕动症缓解方法。因此，我们认为，所提出的基于 XAI 的晕动症模型缩减方法及其部署模型，有助于研究人员在一体式虚拟现实头显上实时自动检测晕动症，并提升虚拟现实的可用性。

7 局限性与未来工作

尽管我们提出的、基于可解释人工智能的超级学习器模型在晕动症检测、特征降维和部署方面优于最先进的晕动症检测模型，但我们的方法仍有一些局限性。例如，我们展示了所提出的基于可解释人工智能的模型缩减方法的有效性，该方法具有较快的训练和推理速度；我们还将模型部署在一台基于高通骁龙处理器的三星 A52 设备上（最先进的虚拟现实头戴式显示器使用高通骁龙处理器）。然而，我们并未将模型部署到实际的虚拟现实头显上。因此，很难说明哪种传感器适用于晕动症检测。不过，根据我们提出的基于可解释人工智能的方法，眼动追踪传感器在晕动症预测方面远比头部追踪传感器高效。因此，值得一提的是，外部传感器（例如心率、皮肤电反应和脑电图传感器）可能会限制沉浸过程中的虚拟现实移动和三维物体操控。此外，这些传感器通常需要通过线缆连接，并固定在用户的手上。

此外，晕动症可能会因不同个体的独特特征、虚拟现实环境和性别 [86]而对他们产生不同影响。例如，瞳距（IPD）无法与虚拟现实头显正确匹配的女性，往往会出现严重的晕动症，且短时间内无法完全恢复 [28, 86]。因此，未来我们计划针对人口背景更广泛、性别比例均衡的人群开展进一步研究。此外，本研究仅使用眼动追踪、头部追踪、生理信号和用户个人资料数据来检测晕动症。未来，我们计划研究立体图像和立体视频数据在可解释晕动症检测中的作用，并将其部署到逼真的虚拟现实头显中。

8 结论

在本研究中，我们提出了 VR-LENS 框架，这是一个基于可解释人工智能的晕动症检测框架，通过基于超级学习器的集成机器学习模型提供解释，并将其部署于嵌入式设备。具体而言，我们开发了一个用于晕动症检测的基于超级学习器的集成机器学习模型。随后，我们应用可解释人工智能来解释晕动症，并缩减特征空间和模型大小，以便将模型部署到嵌入式设备上。我们使用三个数据集展示了所提方法的有效性，即集成传感器数据集、生理数据集和游戏数据集。全局解释结果表明，在集成传感器数据集中，眼动追踪特征是导致晕动症的最重要因素。因此，对于生理数据集，皮肤电反应和心率是最重要的特征；对于游戏数据集，玩家位置和用户眼镜则是导致晕动症的最重要特征。此外，我们通过局部解释，为每个样本（误分类实例）发现了更多有价值的见解。最后，基于可解释人工智能的特征排序，我们显著缩减了超级学习器模型的大小，并将其部署在基于高通骁龙处理器的三星 A52 设备上。部署后的超级学习器模型显著缩短了训练时间（最多缩短 1.91 倍）和推理时间（最多缩短 2.46 倍）。例如，对于集成传感器数据集，我们部署的精简超级学习器模型在晕动症分类和回归任务中的准确率为 96%，均方根误差为 0.03，优于现有最先进的研究。据我们所知，这是首项应用可解释人工智能解释基于超级学习器的集成模型、缩减模型大小并将其部署到嵌入式设备上的研究。我们相信，本研究将有助于未来从事晕动症检测、缓解以及一体式虚拟现实头显中晕动症实时预测的研究人员。

致谢

本材料基于美国国家科学基金会通过资助编号 CNS-2114035 提供支持的研究工作。本文中表达的任何观点、findings、结论或建议均为作者所有，并不一定反映美国国家科学基金会的观点。

参考文献

[1] 2021. 用于分类虚拟现实沉浸所致晕动症的机器—深度—集成学习模型。网络心理学、行为与社会网络 24, 11 (2021), 729–736.[2] Kunjal Ahir、Kajal Govani、Rutvik Gajera 和 Manan Shah。2020年。虚拟现实在提升教育学习、军事训练和体育方面的应用。增强人类研究 5, 1 (2020), 1–9.[3] Monagi H Alkinani 和 Mahmoud R El-Sakka。2017年。基于图像块的模型与图像去噪算法：用于降低加性噪声的基于图像块图像去噪方法比较综述。*EURASIP*图像与视频处理期刊 2017, 1 (2017), 1–27.[4] Naomi S Altman。1992年。核与最近邻非参数回归导论。《美国统计学家》 46, 3 (1992), 175–185.[5] 白俊杰、陆芳、张科等。2019年。ONNX：开放神经网络交换格式。*GitHub* 仓库 (2019).[6] Yoshua Bengio 和 Yves Grandvalet。2003年。k折交叉验证的方差不存在无偏估计量。《神经信息处理系统进展》 16 (2003)。

[7] Lisa Castaneda 和 Marc Pacampara。2016年。课堂中的虚拟现实——对硬件、管理、内容和教学法的探索。载于信息技术与教师教育国际会议。教育计算促进协会（AACE），527–534.[8] Tsz Tai Chan、Yixuan Wang、Richard Hau Yue So 和 Jerry Jia。2022年。预测VR头显中镜头畸变在VR场景前庭眼反射期间引起的主观不适。IEEE可视化与计算机图形学汇刊 (2022), 1–1。https: //doi.org/10.1109/TVCG.2022.3168190[9] Tsz Tai Chan、Yixuan Wang、Richard Hau Yue So 和 Jerry Jia。2022年。预测VR头显中镜头畸变在VR场景前庭眼反射期间引起的主观不适。IEEE可视化与计算机图形学汇刊 (2022).[10] Umer Asghar Chattha、Uzair Iqbal Janjua、Fozia Anwar、Tahir Mustafa Madni、Muhammad Faisal Cheema 和 Sana Iqbal Janjua。2020年。虚拟现实中的晕动病：一项实证评估。IEEE Access 8 (2020), 130486–130499.[11] Tianqi Chen 和 Carlos Guestrin。2016年。XGBoost：可扩展的树提升系统。载于第22届ACM SIGKDD知识发现与数据挖掘国际会议论文集（美国加利福尼亚州旧金山）（KDD '16）。ACM，美国纽约州纽约市，785–794。https: //doi.org/10.1145/2939672.2939785[12] Corinna Cortes 和 Vladimir Vapnik。1995年。支持向量网络。《机器学习》20, 3 (1995), 273–297.[13] David R Cox。1958年。二元序列的回归分析。《英国皇家统计学会杂志：B辑（方法论）》20, 2 (1958), 215–232.[14] Mark S Dennison、A Zachary Wisti 和 Michael D'Zmura。2016年。利用生理信号预测晕动症。《显示器》44 (2016), 42–52.[15] Jiong Dong、Kaoru Ota 和 Mianxiong Dong。2022年。VR游戏为何会引发晕动症？一项关于采集和分析VR游戏头部运动数据集的实证研究。IEEE MultiMedia (2022), 1–1。https://doi.org/10.1109/MMUL.2022.3176142[16] Jiong Dong、Kaoru Ota 和 Mianxiong Dong。2022年。VR游戏为何会引发晕动症？一项关于采集和分析VR游戏头部运动数据集的实证研究。IEEE MultiMedia (2022).[17] Hana R Esmaeel。2015年。应用 Android Studio（SDK）工具。《国际计算机科学与软件工程高级研究期刊》5, 5 (2015).[18] Jerome H Friedman。2001年。贪心函数逼近：梯度提升机。《统计年鉴》(2001), 1189–1232.[19] Zan Gao、Jung Eun Lee、Daniel J McDonough 和 Callie Albers。2020年。COVID-19 疫情期间，虚拟现实运动作为促进老年人健康与福祉的应对策略。, 1986页。[20] Augusto Garcia-Agundez、Christian Reuter、Hagen Becker、Robert Konrad、Polona Caserman、André Miede 和 Stefan Göbel。2019年。开发一种用于确定虚拟现实环境中导致晕动症因素的分类器。fi《健康游戏期刊》8, 6 (2019), 439–444.[21] Luis Eduardo Garrido、Maite Frías-Hiciano、Mariano Moreno-Jiménez、Gabriella Nicole Cruz、Zoilo Emilio García-Batista、Kiero Guerra-Peña 和 Leonardo Adrián Medrano。2022年。聚焦晕动症：普遍性、潜在轨迹、易感性及其对虚拟现实体验的影响。ff《虚拟现实》(2022), 1–25.[22] Eugenia Garza、George S Bailey、Danilo Arruda、Thomas A Stoffregen 等。2022年。头戴式显示器中的头部倾斜及其对晕动症发生率的影响。ff(2022).[23] Pierre Geurts、Damien Ernst 和 Louis Wehenkel。2006年。极端随机树。《机器学习》63, 1 (2006), 3–42.[24] Brandon M Greenwell、Bradley C Boehmke 和 Andrew J McCarthy。2018年。一种简单有效的基于模型的变量重要性度量。arXiv预印本 arXiv:1805.04755 (2018).[25] Aniket Gulhane、Akhil Vyas、Reshmi Mitra、Roland Oruche、Gabriela Hoefer、Samaikya Valluripally、Prasad Calyam 和 Khaza Anuarul Hoque。2019年。虚拟现实学习环境应用的安全性、隐私与安全风险评估。载于2019年第16届IEEE消费者通信与网络年会（CCNC）。IEEE，1–9.[26] Azadeh Hadadi、Christophe Guillet、Jean-Rémy Chardonnet、Mikhail Langovoy、Yuyang Wang 和 Jivka Ovtcharova。2022年。利用拓扑数据分析和机器学习预测虚拟环境中的晕动症。Frontiers in Virtual Reality 3 (2022).[27] A Halbig、SK Babu、S Gatter、ME Latoschik、K Brukamp 和 S von Mammen。2022年。虚拟现实在医疗保健领域的机遇与挑战——领域专家调查。Front. Virtual Real. 3: 837616. doi: 10.3389/frvir (2022).[28] Ancella Hendrika、Clara Theresia 和 Thedy Yogasara。2020年。利用虚拟现实测试性别与经验因素对晕动症的影响。《国际工程技术与自然科学期刊》2, 2 (2020), 63–69.[29] Matt C Howard 和 Elise C Van Zandt。2021年。虚拟现实问题的元分析：虚拟现实晕动症对个体差异的影响并不均等。ff《虚拟现实》25, 4 (2021), 1221–1246。[30] samsung。Inc. 2022。三星5G设备。https://www.samsung.com/us/smartphones/galaxy-a52-5g/[31] Rifatul Islam。2020年。一种基于深度学习的框架，用于检测并减轻晕动症的发生。载于2020年IEEE虚拟现实与三维用户界面会议摘要与研讨会（VRW）。IEEE，559–560.[32] Rifatul Islam、Samuel Ang 和 John Quarles。2021年。Cybersense：用于检测晕动症严重程度并自适应应用缓解技术的闭环框架。载于2021年IEEE虚拟现实与三维用户界面会议摘要与研讨会（VRW）。IEEE，148–155.[33] Rifatul Islam、Kevin Desai 和 John Quarles。2021年。基于头戴式显示器传感器融合的晕动症预测：一种结合眼动追踪和头部追踪数据的多模态深度融合方法。载于2021年IEEE国际混合与增强现实研讨会（ISMAR）。IEEE，31–40.[34] Rifatul Islam、Kevin Desai 和 John Quarles。2022年。通过多模态深度融合网络融合生理信号、头部追踪和眼动追踪数据，以期预测晕动症的发生。

[35] Rifatul Islam、Yonggun Lee、Mehrad Jaloli、Imtiaz Muhammad、Dakai Zhu、Paul Rad、Yufei Huang 和 John Quarles。2020年。利用用户生理信号和深度神经网络自动检测和预测晕动症严重程度。载于2020年IEEE国际混合与增强现实研讨会（ISMAR）。IEEE，400–411。[36] Rifatul Islam、Yonggun Lee、Mehrad Jaloli、Imtiaz Muhammad、Dakai Zhu、Paul Rad、Yufei Huang 和 John Quarles。2020年。利用用户生理信号和深度神经网络自动检测和预测晕动症严重程度。载于2020年IEEE国际混合与增强现实研讨会（ISMAR）。400–411。https://doi.org/10.1109/ISMAR50242.2020.00066[37] Dayoung Jeong 和 Kyungsik Han。2022年。在晕动症预测中利用多模态感官信息。载于第28届ACM虚拟现实软件与技术研讨会论文集。1–2。[38] Daekyo Jeong、Sangbong Yoo 和 Jang Yun。2019年。使用深度学习算法结合脑电图进行晕动症分析。载于2019年IEEE虚拟现实与三维用户界面会议（VR）。IEEE，827–835。[39] D. Jeong、S. Yoo 和 J. Yun。2019年。使用深度学习算法结合脑电图进行晕动症分析。载于2019年IEEE虚拟现实与三维用户界面会议（VR）。827–835。https://doi.org/10.1109/VR.2019.8798334[40] Weikuan Jia、Meili Sun、Jian Lian 和 Sujuan Hou。2022年。特征降维：综述。《复杂与智能系统》（2022），1–31。[41] Weina Jin、Jianyu Fan、Diane Gromala 和 Philippe Pasquier。2018年。虚拟现实游戏中的晕动症自动预测。载于2018年IEEE游戏、娱乐与媒体会议（GEM）。IEEE，1–9。[42] Robert S Kennedy、Norman E Lane、Kevin S Berbaum 和 Michael G Lilienthal。1993年。模拟器晕动症问卷：一种量化模拟器晕动症的改进方法。《国际航空心理学杂志》3，3（1993），203–220。[43] Behrang Keshavarz 和 Heiko Hecht。2011年。验证一种量化晕动病的高效方法。《人为因素》53，4（2011），415–426。fiBehrang Keshavarz、Katlyn Peck、Sia Rezaei 和 Babak Taati。2022年。结合生理测量指标与机器学习技术检测和预测视觉诱发晕动病。《国际心理生理学杂志》（2022）。[44] Hyun K Kim、Jaehyun Park、Yeongcheol Choi 和 Mungyeong Choe。2018年。虚拟现实病症问卷（VRSQ）：虚拟现实环境中的晕动病测量指标。《应用工效学》69（2018），66–73。[45] Jinwoo Kim、Woojae Kim、Heeseok Oh、Seongmin Lee 和 Sanghoon Lee。2019年。基于脑信号分析的虚拟现实内容深度晕动症预测器。载于IEEE/CVF国际计算机视觉会议论文集。10580–10589。[46] Young Youn Kim、Eun Nam Kim、Min Jae Park、Kwang Suk Park、Hee Dong Ko 和 Hyun Taek Kim。2008年。利用生物信号反馈减少暴露于虚拟环境所引起的晕动症。《临场感：远程操作与虚拟环境》17，1（2008），1–16。[47] R Kottaimalai、M Pallikonda Rajasekaran、V Selvam 和 B Kannapiran。2013年。在脑机接口应用中使用结合主成分分析与神经网络的脑电信号分类。载于2013年IEEE计算、通信与纳米技术新兴趋势国际会议（ICECCN）。IEEE，227–231。[48] Ripan Kumar Kundu、Rifatul Islam、Prasad Calyam 和 Khaza Anuarul Hoque。2022年。TruVR：使用可解释机器学习进行可信晕动症检测。arXiv预印本 arXiv:2209.05257（2022）。fiRipan Kumar Kundu、Akhlaqur Rahman 和 Shuva Paul。2021年。VR晕动病中的传感器系统延迟研究。《传感器与执行器网络期刊》10，3（2021），53。[49] Joseph J LaViola Jr. 2000年。虚拟环境中的晕动症讨论。《ACM SIGCHI通报》32，1（2000），47–56。[50] Tae Min Lee、Jong-Chul Yoon 和 In-Kwon Lee。2019年。使用三维卷积神经网络预测立体视频中的晕动病。《IEEE可视化与计算机图形学汇刊》25，5（2019），1919–1927。[51] Ajey Lele。2013年。虚拟现实及其军事用途。《环境智能与人性化计算期刊》4，1（2013），17–26。[52] Andy Liaw 和 Matthew Wiener。2002年。使用 randomForest 进行分类与回归。《R新闻》2，3（2002），18–22。https://CRAN.R-project.org/ doc/Rnews/[53] Chin-Teng Lin、Shu-Fang Tsai 和 Li-Wei Ko。2013年。用于动态车辆环境中在线估计晕动病程度的基于脑电图的学习系统。《IEEE神经网络与学习系统汇刊》24，10（2013），1689–1700。[54] Yi-Tien Lin、Yu-Yi Chien、Hsiao-Han Wang、Fang-Cheng Lin 和 Yi-Pai Huang。2018年。65-3：利用脑电图和心电图量化虚拟现实头戴式显示器中的晕动症等级。载于《SID技术论文研讨会文摘》，第49卷。Wiley在线图书馆，862–865。fiScott M Lundberg 和 Su-In Lee。2017年。解释模型预测的统一方法。载于《神经信息处理系统进展》第30卷，I. Guyon、U. V. Luxburg、S. Bengio、H. Wallach、R. Fergus、S. Vishwanathan 和 R. Garnett（编）。Curran Associates, Inc.，4765–4774。http://papers.nips.cc/paper/7062-a-uni[55] ed-approach-to-interpreting-model-predictions.pdf[56] Måns Magnusson、Michael Andersen、Johan Jonasson 和 Aki Vehtari。2019年。针对大规模数据的贝叶斯留一交叉验证。载于国际机器学习会议。PMLR，4244–4253。[57] Sannia Mareta、Joseph Manuel Thenara、Rafael Rivero 和 May Tan-Mullins。2022年。虚拟现实网络病影响及面向本科生整体虚拟学习体验的改进策略研究。《交互技术与智慧教育》即将刊载（2022）。fiNicolas Martin、Nicolas Mathieu、Nico Pallamin、Martin Ragot 和 Jean-Marc Diverrez。2020年。虚拟现实病症检测：一种基于生理信号和机器学习的方法。载于2020年IEEE国际混合与增强现实研讨会（ISMAR）。IEEE，387–399。fiMoch Asyroful Mawalid、Al[58] Zuhriya Khoirunnisa、Mauridhi Hery Purnomo 和 Adhi Dharma Wibawa。2018年。使用朴素贝叶斯通过时域特征提取对脑电信号进行分类以检测晕动症。载于2018年计算机工程、网络与智能多媒体国际会议（CENIM）。IEEE，29–34。

[62] Miguel Melo、José Vasconcelos-Raposo 和 Maximino Bessa。2018年。沉浸式内容中的临场感与晕动症：内容类型、暴露时间和性别的影响。《Computers & Graphics》71（2018），159–165。ffZahira Merchant、Ernest T Goetz、Lauren Cifuentes、Wendy Keeney-Kennicutt 和 Trina J Davis。2014年。基于虚拟现实的教学对K-12 教育和高等教育学生学习成果的有效性：元分析。《Computers & Education》70（2014），29–40。[63] Max D. Morris。1991年。初步计算实验的析因抽样方案。《Technometrics》33，2（1991年5月），161–174。https: //doi.org/10.1080/00401706.1991.10484804ffSyed Ali Arsalan Naqvi、Nasreen Badruddin、Munsif Ali Jatoi、Aamir Saeed Malik、Wan Hazabbah 和 Baharudin Abdullah。2015年。视觉诱发晕动病（VIMS）的脑电图时频动态分析。《Australasian physical & engineering sciences in medicine》38，4（2015），721–729。[64] Harsha Nori、Samuel Jenkins、Paul Koch 和 Rich Caruana。2019年。InterpretML：机器学习可解释性的统一框架。arXiv预印本 arXiv:1909.09223（2019）。[65] Nitish Padmanaban、Timon Ruban、Vincent Sitzmann、Anthony M Norcia 和 Gordon Wetzstein。2018年。面向 360 度立体视频晕动症预测的机器学习方法。《IEEE可视化与计算机图形学汇刊》24，4（2018），1594–1603。[66] Amirhossein Panahi、Reza Askari Moghadam、Bahram Tarvirdizadeh 和 Kurosh Madani。2022年。用于青光眼检测的简化U-Net深度学习智能医疗辅助工具。《Evolutionary Intelligence》（2022），1–12。fiEvi Septiana Pane、Al[67] Zuhriya Khoirunnisaa、Adhi Dharma Wibawa 和 Mauridhi Hery Purnomo。2018年。使用CN2规则归纳算法从脑电信号中识别晕动症严重程度等级。载于2018年智能信息学与生物医学科学国际会议（ICIIBMS），第3卷。170–176。https://doi.org/10.1109/ICIIBMS.2018.8549968[68] Su Han Park、Bin Han 和 Gerard Jounghyun Kim。2022年。在虚拟现实中混合反向光流以减轻自运动感和模拟器晕动症。载于CHI计算机系统中的人为因素会议。1–11。fiFabian Pedregosa、Gaël Varoquaux、Alexandre Gramfort、Vincent Michel、Bertrand Thirion、Olivier Grisel、Mathieu Blondel、Peter Prettenhofer、Ron Weiss、Vincent Dubourg 等。2011年。Scikit-learn：Python中的机器学习。《Journal of Machine Learning Research》12（2011），2825–2830。[69] F. Pedregosa、G. Varoquaux、A. Gramfort、V. Michel、B. Thirion、O. Grisel、M. Blondel、P. Prettenhofer、R. Weiss、V. Dubourg、J. Vanderplas、A. Passos、D. Cournapeau、M. Brucher、M. Perrot 和 E. Duchesnay。2011年。Scikit-learn：Python中的机器学习。《Journal of Machine Learning Research》12（2011），2825–2830。fiThiago Porcino、Erick O Rodrigues、Flavia Bernardini、Daniela Trevisan 和 Esteban Clua。2021年。使用符号机器学习算法识别虚拟现实游戏中的晕动症成因。《Entertainment Computing》（2021），100473。[70] Thiago Porcino、Erick O. Rodrigues、Alexandre Silva、Esteban Clua 和 Daniela Trevisan。2020年。利用游戏玩法和用户数据预测并识别虚拟现实游戏中晕动症表现的成因。载于2020年IEEE第八届严肃游戏与健康应用会议（SeGAH）。1–8。 flChenxin Qu、Xiaoping Che、Siqi Ma 和 Shuqin Zhu。2022 年。基于生物生理信号的 VR 晕动症检测。《 CCF Transactions on Pervasive Computing and Interaction 》（2022），1–17。[71] Lisa Rebenitsch 和 Charles Owen。2016年。应用和视觉显示设备中的晕动症综述。《虚拟现实》20，2（2016），101–125。[72] Marco Recenti、Carlo Ricciardi、Romain Aubonnet、Ilaria Picone、Deborah Jacob、Halldór ÁR Svansson、Sólveig Agnarsdóttir、Gunnar H Karlsson、Valdís Baeringsdóttir、Hannes Petersen 等。2021年。利用虚拟现实和移动平台评估大脑、肌肉及心脏信号，探索运动病预测。《Frontiers in Bioengineering and Biotechnology》9（2021），132。[73] GS Reddy、Cara A Spencer、Kevin Durkee、Brennan Cox、Olivia Fox Cotton、Sheila Galbreath、Sarah Meyer、Michael Natali、Todd Seech、Gabriella Severe-Valsaint 等。2022年。通过非侵入式生理传感器估计VR模拟中飞行员的认知负荷和晕动症。载于人机交互国际会议。Springer，251–269。[74] Paula Reyero Lobo 和 Pablo Perez。2022年。利用心率变异性进行非侵入式晕动症检测。载于ACM交互媒体体验国际会议。221–228。[75] Marco Tulio Ribeiro、Sameer Singh 和 Carlos Guestrin。2016年。“我为什么应该相信你？”解释任意分类器的预测。载于第22届ACM SIGKDD知识发现与数据挖掘国际会议论文集。1135–1144。[76] Omer Sagi 和 Lior Rokach。2018年。集成学习：综述。《Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery》8，4（2018），e1249。[77] Md Satu、Shelia Rahman、Md Khan、Mohammad Zoynul Abedin、M Shamim Kaiser、Mufti Mahmud 等。2020年。利用双向多层长短期记忆神经网络改进认知表现检测。载于脑信息学国际会议。Springer，297–306。[78] Erick J Schmitt 和 Hossein Jula。2007年。线性模型在出行时间预测中的局限性。载于2007年IEEE智能交通系统会议。IEEE，830–835。[79] Volkan Sevinc 和 Mehmet Ilker Berkman。2020年。将模拟器晕动症问卷及其变体作为消费者虚拟环境中晕动症测量工具的心理测量学评估。《应用工效学》82（2020），102958。[80] Ashutosh Singla、Steve Göring、Dominik Keller、Rakesh Rao Ramachandra Rao、Stephan Fremerey 和 Alexander Raake。2021年。全向视频的模拟器晕动症问卷评估。载于2021年IEEE虚拟现实与三维用户界面会议（VR）。198–206。https: // doi.org/10.1109/VR50410.2021.00041fiKay Stanney、Cali Fidopiastis 和 Linda Foster。2020年。虚拟现实具有性别歧视：但不必如此。《Frontiers in Robotics and AI》7（2020），4。

[87] Kay M Stanney、Robert S Kennedy 和 Julie M Drexler。1997年。晕动症并非模拟器晕动症。载于人因与工效学学会年度会议论文集，第41卷。SAGE Publications Sage CA：洛杉矶，加利福尼亚州，1138–1142。[88] Scott Stein。2020年。Zoom，不过是在虚拟现实中：为什么 spatial 的免费会议应用让人感觉实现了飞跃。https://www.cnet.com/tech/gaming/zoom-but-in-vr-why-spatials-free-meeting-app-feels-like-a-leap-forward/[89] William Bruce Stone III。2017年。作为晕动症测量工具的《模拟器晕动症问卷》心理测量学评估。博士论文。爱荷华州立大学。[90] Nana Tian、Phil Lopes 和 Ronan Boulic。2022年。头戴式显示器中的晕动症综述：关注个体易感性。虚拟现实（2022），1–33。[91] Erico Tjoa 和 Cuntai Guan。2021年。可解释人工智能（XAI）综述：迈向医疗可解释人工智能。*IEEE*神经网络与学习系统汇刊32，11（2021），4793–4813。https://doi.org/10.1109/TNNLS.2020.3027314[92] Ehsan Toreini、Mhairi Aitken、Kovila Coopamootoo、Karen Elliott、Carlos Gonzalez Zelaya 和 Aad Van Moorsel。2020年。对人工智能的信任与可信机器学习技术之间的关系。载于*2020*年公平性、问责制与透明度会议论文集。272–283。[93] Samaikya Valluripally、Benjamin Frailey、Brady Kruse、Boonakij Palipatana、Roland Oruche、Aniket Gulhane、Khaza Anuarul Hoque 和 Prasad Calyam。2022年。检测干扰用户在虚拟现实学习环境中沉浸式体验的安全与隐私攻击。*IEEE*服务计算汇刊（2022）。[94] Samaikya Valluripally、Aniket Gulhane、Khaza Anuarul Hoque 和 Prasad Calyam。2021年。社交虚拟现实攻击诱发晕动症的建模与防御。*IEEE*可信与安全计算汇刊（2021）。[95] Harshita Ved 和 Caglar Yildirim。2021年。利用脑电频谱数据检测虚拟现实中的心理工作负荷：一种深度学习方法。载于*2021*年*IEEE*人工智能*fi*与虚拟现实国际会议（*AIVR*）。IEEE，173–178。[96] Jialin Wang、Rongkai Shi、Zehui Xiao、Xueying Qin 和 Hai-Ning Liang。2022年。虚拟现实游戏中游戏体验、性能与模拟器晕动症之间的分辨率权衡。载于*2022*年*IEEE*虚拟现实与三维用户界面会议摘要与研讨会（*VRW*）。542–543。https://doi.org/10.1109/VRW55335.2022.00122[97] Yuyang Wang、Jean-Rémy Chardonnet、Frédéric Merienne 和 Jivka Ovtcharova。2021年。利用模糊逻辑纳入个体差异，预测虚拟现实导航过程中的晕动症。载于*2021*年*IEEE* 虚拟现实与三维用户界面会议（*VR*）。IEEE，373–381。[98] Brenda K Wiederhold。2020年。2019冠状病毒病疫情期间通过技术保持联系：避免"Zoom疲劳"。[99] Alexander Hui Xiang Yang、Nikola Kasabov 和 Yusuf Ozgur Cakmak。2022年。用于晕动症研究的机器学习方法：系统综述。*BrainInformatics* 9，1（2022），1–25。[100] Yoichi Yokota、Mitsuhiro Aoki、Keisuke Mizuta、Yatsuji Ito 和 Naoki Isu。2005年。健康受试者中视觉诱发的姿势不稳定与心脏自主神经反应相关的晕动症易感性。*Acta oto-laryngologica* 125，3（2005），280–285。[101] Zhi-Hua Zhou。2012年。集成方法：基础与算法（第0版）。Chapman and Hall/CRC。https://doi.org/10.1201/b12207[102] Kelly H Zou、Kemal Tuncali 和 Stuart G Silverman。2003年。相关与简单线性回归。*Radiology* 227，3（2003），617–628。