arXiv:2602.11668v1 [cs.LG] 12 Feb 2026

基于可解释机器学习的跑步者膝关节损伤检测

预印本

David Fuentes-Jiménez[*1]、**Sara García-de-Villa**[†2]、**David Casillas-Pérez**[‡2]、**Pablo Floría**[§3]和 **Francisco-Manuel Melgarejo-Meseguer**[¶2]

[1]阿尔卡拉大学电子学系，西班牙阿尔卡拉-德埃纳雷斯，28801 [2]胡安·卡洛斯国王大学信号处理系，西班牙丰拉夫拉达，28942[3]巴勃罗·德奥拉维德大学体能与体育研究中心 (CIRFD)，西班牙塞维利亚，41013

2026年2月13日

摘要

跑步是一项广泛开展的休闲活动，但其膝关节损伤发生率较高，尤其是髌股疼痛综合征（PFPS）和髂胫束综合征（ITBS）。准确识别与这些损伤相关的模式，有望影响临床决策。实现这一目标需要能够精准捕捉并分析人体运动时序运动学数据的系统。在本研究中，采用光学动作捕捉系统改进损伤相关步态模式的检测。研究分析了一个公开数据集，其中包含来自健康跑步者和受伤跑步者的839条跑步机运动记录，以评估这些测量系统在捕捉与损伤分类相关的动态运动参数方面的有效性。本研究聚焦于跑步步态的支撑期，考察从动作捕捉数据中获得的关节角度和肢段角度时间序列，以及离散点值。我们评估了三种分类场景：（1）健康跑步者与受伤跑步者；（2）健康跑步者与PFPS患者；（3）健康跑步者与ITBS患者。我们探索了多种特征输入空间，从传统的基于点的指标到支撑期的完整时间序列表征，以及结合二者的混合方法。研究采用了多种分类模型，包括K近邻、高斯过程和决策树等经典算法，以及卷积神经网络（CNN）和长短期记忆网络（LSTM）等深度学习架构。模型性能采用准确率、精确率、召回率和F1分数进行评估。此外，我们还结合了Shapley值、部分依赖图、显著性图和Grad-CAM等可解释性技术，以解读模型的决策。结果表明，将时间序列数据与点值相结合能够显著提升检测性能。深度学习模型优于传统方法，其中卷积神经网络取得了最高准确率：PFPS为77.9%，ITBS为73.8%，合并损伤类别（PFPS+ITBS）为71.43%。这些发现凸显了结合先进机器学习技术的动作捕捉系统在识别膝关节损伤相关跑步模式方面的有效性。

资助： 本研究部分由西班牙科学、创新与大学部（MICIU/AEI/10.13039/501100011033）、HERMES项目（资助编号：PID2023-152331OA-I00）、METAMORPH项目（资助编号：PID2023-151295OB-I00）和LATENTIA项目（资助编号：PID2022-140786NB-C31）资助，并获得马德里自治区CAM 2024产业博士项目（IND2024/TIC-33917）的资助。

[*]d.fuentes@uah.es
[†]sara.garcia.devilla@urjc.es
[‡]david.casillas@urjc.es[§]
pfloriam@upo.es
[¶]francisco.melgarejo@urjc.es

***Key*词语** 运动分析 ·跑步模式 · 髌股 · 髂胫束 · 机器学习

1 引言

跑步是最受欢迎的休闲体育活动之一。然而，跑步也伴随着较高的肌肉骨骼损伤发生率，每跑步 1000 小时会发生 4 至 33 例损伤 Videbaek 等人 [2015]。在休闲跑者中，膝关节是最常受影响的关节，最常见的诊断是髂胫束综合征（ITBS）和髌股疼痛综合征（PFPS）Taunton 等人 [2002]。这些损伤的成因多种多样，涉及外部变量（例如训练负荷、运动量）与内部因素（例如生物力学、年龄）之间的复杂相互作用 Van der Worp 等人 [2015]。

识别与健康状况相关的运动模式，有助于制定预后指标，从而通过步态干预支持损伤预防和长期健康维护 Crowell 和 Davis[2011], Soltanabadi 等人 [2023], Keast 等人 [2022]。对于跑步者而言，检测此类模式需要准确量化跑步过程中的运动，随后进行单参数或多参数分析，并应用数据分析算法。

既往研究表明，步态模式分析在预测不良健康结局方面具有潜力，例如神经退行性疾病（Cicirelli 等人 [2021a], Hammoud 等人 [2025], Shcherbak 等人 [2023] ）或损伤风险（Lee 等人 [2022], Zeng 等人 [2020], Kokkotis 等人 [2022], ），所采用的技术包括基于摄像机的系统和可穿戴传感器技术。这些发现凸显了运动分析在临床和预防应用中的价值。

尽管步态分析取得了进展，但跑步中的损伤相关运动模式研究仍是一个新兴领域。多项研究使用光学动作捕捉系统，尝试识别受伤跑步者之间相似的运动学步态模式（Zhuang 等人 [2025], ），但目前的证据尚不足以建立有意义的跑步相关风险特征（Ceyssens 等人 [2019]）。这种结论不一致可能源于样本量有限、将时间序列数据简化为生物力学离散值（如峰值）时造成的信息损失（Phinyomark 等人 [2017], ），或对跑步相关损伤进行笼统研究（Willwacher 等人 [2022], ），从而夸大了特定生物力学风险因素对具体损伤的重要性（Ceyssens 等人 [2019]）。

模式识别技术（如聚类算法）被用于对跑步者进行分类，并揭示与损伤风险相关的生物力学模式 Dingnenen 等人 [2020], Jauhiainen 等人 [2020], Martin 等人 [2022], Senevirathna 等人[2023]。然而，结果表明，虽然可以识别出运动模式相似的亚组，但尚未发现这些亚组与损伤风险之间存在直接关联。据我们所知，目前仅有一种监督式机器学习（ML）方法被用于损伤风险检测 Eskofier 等人 [2012], ，具体而言，是用于识别髌股疼痛综合征（PFPS）。因此，其他监督式机器学习方法尚未用于此类分析。机器学习方法使用有限的一个可能原因是其黑箱行为，这限制了它们在临床应用中的使用。然而，可解释性算法的最新进展使得揭示模型预测与其输入之间的关系成为可能。

本研究旨在探讨光学动作捕捉系统提供的关节角度时间序列和动力学描述符在区分膝关节受伤跑者与健康跑者方面的有效性，并评估监督式机器学习方法的性能。所分析的机器学习方法包括经典和现代算法，例如深度学习（DL）模型。此外，我们还通过应用可解释性算法揭示所用方法的黑箱行为，从而提高研究结果的可解释性。
本文在 Fuentes-Jiménez 等人 [2025], 发表于会议论文集的研究基础上，扩展了所分析的深度学习算法，并增强了方法的可解释性。本研究现纳入卷积神经网络（CNN）和长短期记忆网络（LSTM），以及经典机器学习模型。文中给出了所有算法的分类指标，并详细介绍了超参数优化过程。此外，通过纳入四种新方法，进一步扩展了可解释性分析，从而为性能最佳的算法提供了更深入的见解。

本文其余部分的结构如下：第 2 节回顾使用运动学信息进行损伤模式识别的相关研究；第 3 节介绍本研究所使用的跑者数据库、所采用的方法论（包括所使用的算法、算法的实现及其分析指标）以及可解释性方法；第 4 节呈现所分析分类器的分类指标和可解释性分析结果，并深入考察特征，随后讨论这些结果；最后，第 5 节总结使用监督算法检测损伤模式以及基于可解释性进行特征分析的结论性观点。

2 研究现状

运动时间动态的研究已得到广泛开展，尤其是在步态研究领域，主要用于检测神经系统疾病 Kokkotis 等人 [2022], Cicirelli 等人 [2021b]。例如，在 Xia 和 Cao [2026], Raja 等人 [2026] 中，从视频记录和可穿戴传感器中提取人体运动动力学信息，随后通过图卷积网络进行处理，以检测帕金森病患者的步态冻结。在 Ru 等人 [2025]a 中，研究人员使用步态分析，并结合机器学习方法，对帕金森病、肌萎缩侧索硬化症和亨廷顿病进行分类。
关于运动相关损伤，Moura 等人 [2025], 采用步态运动学，通过深度学习方法检测膝骨关节炎。在 García-de-Villa 等人 [2024], 的研究中，则利用惯性分析得出的运动学和步态时空参数评估跌倒风险。

在运动科学领域，采用动态运动分析来降低损伤风险、提升运动表现的研究较少。在 Liu 等人 [2025], 的研究中，训练了一维卷积神经网络，以对训练期间的游泳模式进行分类，从而提高效率。在 Chen 和 Jia [2025], 的研究中，则利用动力链理论分析检测并分类网球击球动作中的复杂动态模式，以提高运动效率并预防运动损伤。

聚焦于跑步损伤，Eskofier 等人 [2012] 的研究是最早采用通用生物医学特征对性别、穿鞋／赤足状态及损伤模式组进行分类的研究之一，具体关注髌股疼痛综合征与未受伤跑者的差异。该研究使用 AdaBoost 进行分类，并通过留一被试交叉验证进行测试。在匹配未受伤者与受伤者的跑步里程、经验和体重后，研究依据髋关节外展力矩完美识别出髌股疼痛综合征。上述结果有望用于识别髌股疼痛综合征与未受伤者跑步模式之间的差异。然而，研究结果仅基于六名受伤志愿者；对更大规模数据库的进一步分析也未能确定一种简单的髌股疼痛综合征模式识别方法。

近期，Dingenen 等人 [2020] 对53名患有跑步相关损伤的休闲跑者开展了规模更大的分析。该研究评估了初始接触时的足部和胫骨倾斜角，以及支撑中期额状面和矢状面上的髋关节内收与膝关节屈曲；这些指标通过视频标记分析，在跑步者以自选速度穿鞋在跑步机上跑步时进行追踪。研究采用 K均值聚类分析对四项结局指标进行聚类，并使用轮廓系数法确定最佳聚类数。该研究未能确定损伤的任何模式：同一种损伤可能对应不止一种运动学表征，而一种模式也可能与不同损伤相关。

Jauhiainen 等人 [2020] 的研究在一大群受伤和未受伤跑者中识别出具有同质跑步模式的不同亚组，共有291名受伤和未受伤跑者，涵盖不同性别和广泛的年龄范围（$10-66$ 岁）。研究根据动作捕捉提供的运动学数据，按损伤部位类别（膝关节、踝部/足部和髋部/骨盆）采用层次聚类方法计算聚类结果。研究结果表明，具有相同损伤类型的跑步者并不会聚在一起。
层次聚类也被用于根据步态动力学模式，在一个包含134名受伤和未受伤跑步者的跑者数据库中检测不同聚类 Senevirathna 等人 [2023]。步态模式依据通常与跑步损伤相关的步态参数进行分析，例如触地时间、平均加载率或垂直冲量。结果揭示了跑步者的聚类，但这些聚类与损伤状态均无关联。

层次聚类也被应用于特定损伤，如 Martin 等人 [2022], 将其用于识别骨应力性损伤（BSI）风险。本研究利用来自双能*X*射线吸收测定法 （DXA）扫描的骨盆和股骨几何结构，分析了一个包含53名大学跑者的数据库。研究中通过主成分分析对几何结构进行降维。结果给出七个不同组别，揭示各组在骨盆和股骨近端几何结构方面的差异，这些差异与随后学年中的下肢骨应力损伤发生率相关。
以往旨在寻找损伤相关模式的研究主要侧重于无监督方法，这些方法主要将速度等瞬时运动学描述符作为点值进行评估。其中大多数使用层次聚类方法对跑步者进行分组 Jauhiainen 等人 [2020], Martin 等人 [2022], Senevirathna 等人 [2023]。所综述的研究均未对监督方法或无监督方法进行比较，以验证所得聚类是否一致。除 Eskofier 等人 [2012], 外，没有将监督方法应用于分类任务，主要原因是难以识别最相关的特征；此外，它们均未依赖近期的深度神经网络（深度神经网络）模型。本文旨在对更多监督式分类技术在跑步损伤风险识别中的应用进行更深入的分析，并重点关注其可解释性。

除了 Martin 等人 [2022], 使用基于 DXA扫描 得出的参数外，大多数研究仅依赖基于点的特征，例如速度、关节倾角、特定时刻的时间值或触地时间，而没有将完整的动力学纳入模型输入。所使用的数据库通常规模较小，这限制了它们对不同类型伤病跑步者群体的

代表性。大多数情况下，所分析的伤病类型并未具体到身体部位，Eskofier 等人 [2012] 的研究例外，该研究聚焦于髌股疼痛综合征。这种跑步模式的异质性凸显了按伤病类型进行具体分析的必要性，以识别更具针对性的运动模式，从而实现更好的区分。

Fuentes-Jiménez 等人 [2025] 的研究首次比较了用于检测膝关节损伤的监督式机器学习方法，其中包括一种近期提出的基于深度神经网络的方法，并提供了用于可解释性分析的显著性图。然而，该研究未探讨利用时间依赖关系的替代方法，例如长短期记忆网络；此外，其可解释性分析仅限于在卷积神经网络上应用单一方法，而本研究则对此有所突破。鉴于研究现状存在这些局限，本文提出一种基于监督学习的框架，将完整支撑期时间序列分析与运动学描述符相结合。本研究的主要贡献如下：*a)*使用时间序列和运动学描述符分析用于识别髂胫束综合征和髌股疼痛综合征损伤模式的经典机器学习方法，并详细说明超参数优化过程；*b)* 改进新提出的深度学习方案，尤其是使用时间序列的方案，包括卷积神经网络和长短期记忆网络；*c)* 应用五种新颖的基于可解释性的算法，揭示黑箱行为，从而提高在损伤模式识别中取得最高分类指标的方法的可理解性。

3 材料与方法

3.1 跑步数据库

我们重点研究 Ferber 等人 [2024] 近期发布的跑步数据库，该数据库记录了未受伤跑者和受伤跑者在跑步机跑步期间的数据。简而言之，该数据库由 1798 名具有不同受伤状况的跑步者组成，并包含通过高速光电红外运动捕捉摄像机（Vicon，英国牛津）测量的跑步机跑步运动记录。

我们的研究重点分析两种常见的膝关节损伤：PFPS 和 ITBS。Ferber 等人 [2024]的数据库包括分别患有 PFPS 和 ITBS 的 137 名和 100 名跑步者，以及 396 名未受伤跑者。部分跑步者有多条记录，因此在为每条记录分配一个样本后，我们分别分析了来自 PFPS 和 ITBS 跑步者的 137 条和 126 条记录，以及来自未受伤跑者的 576 条记录。

每个记录样本包括社会人口统计学变量、人体测量学变量以及 时间序列原始标记点轨迹。这些变量描述了志愿者的年龄、优势腿、性别、身高、体重和跑步相关损伤。标记点轨迹描述了七个解剖结构的运动：双脚、双小腿、双大腿和骨盆。

3.2 数据预处理

我们在数据预处理中使用 Ferber 等人数据库补充材料中的代码 [2024]，其流程图如图 1 所示。我们计算六个关节（踝关节、膝关节和髋关节）以及三个身体节段（足部和骨盆）的 X-Y-Z 卡丹角。对于左脚支撑期，对左侧踝关节、膝关节、髋关节、足部和骨盆的角度进行分段。同样，对于右脚支撑期，也对右侧踝关节、膝关节、髋关节、足部和骨盆的角度进行分段。我们还提取跑步点值。根据 Osis 等人[2014]的方法，使用足部相对于全局坐标系的角度，通过主成分分析检测着地和离地事件，从而将时间序列分割为跑步步幅。我们的分析仅限于每次着地与离地事件之间的跑步支撑期。由于每个步幅的采样数可能不同，我们使用三次插值将采样数统一为 101 个。本文使用的时间序列为分段步幅支撑期内的平均角度以及上包络角和下包络角。

3.3 运动学、时空和功率参数

根据归一化角度时间序列，我们推导出以下跑步时空参数：步宽、步频和步幅、摆动时间和支撑时间的百分比、骨盆下沉角及其活动范围、踝关节处于外翻状态的时间百分比、踝背屈角峰值，以及外翻角和旋转角的峰值与活动范围、膝屈曲角峰值，以及内收、外展和旋转角的峰值与活动范围、髋伸展角峰值，以及内收角和旋转角的峰值与活动范围、足进展角、离地时足跟内侧鞭状运动的活动范围、旋前开始和结束时间，以及踝外翻和旋转、膝外展、内收和旋转、髋内收、外展和旋转、骨盆下沉及垂直振荡的速度峰值。这些参数与登记的描述变量合并，构成点值集。此外，我们还纳入各角度旋转时间序列在 0-1Hz（低频，简称 LF）、1-3Hz（中频，简称 MF）和 3-99Hz（高频，简称 HF）频带内的功率，以捕捉跑步速度方面的差异。

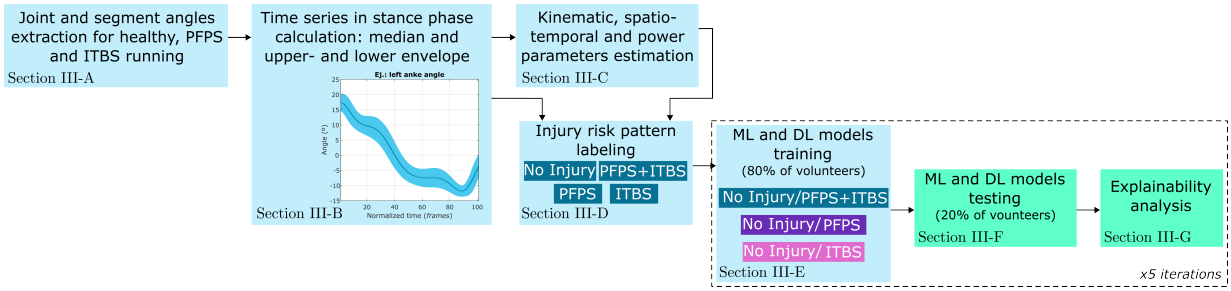


图1：方法流程图展示了数据库中的关节和肢段角度，包括支撑期时间序列的提取和跑步时空参数的计算，直至利用机器学习和深度学习模型识别损伤模式，并对其进行可解释性分析。

3.4 损伤风险识别

我们研究三种不同的二分类：首先，将髌股疼痛综合征（PFPS）和髂胫束综合征（ITBS）合并为一个受伤类别，以检测膝关节受伤跑步者与未受伤跑者；然后，重点检测患有 PFPS 的跑步者与未受伤跑者；最后，仅检测患有 ITBS 的跑步者与未受伤跑者。

我们通过评估模型在不同输入集下的表现，分析以下假设：与点值相比，时间序列包含更多与损伤风险模式检测相关的信息。首先，我们评估分类器仅接收时间序列时的损伤模式识别效果。其次，我们研究将点值添加到时间序列中的影响。最后，为了设定基线情景，我们评估仅以点值作为输入时分类器的表现。

3.5 机器学习分类器

针对损伤风险识别任务评估的分类器包括 8 种监督式机器学习算法和 2 种基于深度学习的方法。下文将介绍所评估的分类器，以及通过网格搜索优化所选的超参数。

K近邻（KNN）Shakhnarovich 等人 [2008],James 等人 [2013] 是一种非参数监督算法。KNN 会在特征空间中找出距离查询点最近的 k 个训练样本，并将这些邻居中最常见的类别标签（即跑步模式类别标签）赋予该查询点。该算法依靠距离度量来确定邻近程度，并使用 k；其取值通过网格搜索进行优化，搜索范围为 2 到 12，步长为 1。

支持向量机（SVM）Schölkopf 等人 [2002], James 等人 [2013] 是一种稳健的监督学习算法。SVM 会寻找最优超平面，使特征空间中不同跑步模式类别之间的间隔最大化。该超平面由支持向量确定，这些支持向量是距离决策边界最近的数据点。核函数使 SVM 无需显式转换数据即可在高维空间中运行。本文使用两种核函数：线性核（SVM_L）和多项式核（SVM_P）。SVM_L 的正则化参数通过网格搜索在以下取值中进行优化：0.0001、0.001、0.005、0.01、0.02、0.03、0.04、0.05 和 0.06。对于 SVM_P，正则化参数通过网格搜索在 0.1 到 1.1 的范围内进行优化，步长为 0.1；并分析了从 2 到 5 的所有次数。

高斯过程（GP）Rasmussen [2003] 是非参数且具有概率性的机器学习模型。高斯过程表示稠密变量的连续随机过程。高斯过程模型通过引入一组来自高斯过程的潜变量来预测跑步模式类别。

决策树（DT）Rokach 和 Maimon [2005], James 等人 [2013] 是基于图的监督式机器学习模型。决策树通过递归地将特征空间划分为互不重叠的不同区域，构建决策过程。决策树根据输入变量的取值划分特征空间，从而预测跑步模式类别。我们对最大深度进行优化，评估不限制最大深度以及最大深度分别为 5、10、20 和 30 的情况。评估的每个节点的样本划分数量为 2、5 和 10，每个叶节点的最小样本数为 1、2 和 4。我们分析基尼准则以进行优化；对于分裂器参数，则评估信息量最大的划分准则和随机划分准则，并通过选择平方根和以2为底的对数来控制特征数量。

自适应提升（ADB） Ferreira 和 Figueiredo [2012],González 等人 [2020] 是一种提升集成机器学习方法。它通过依次组合多个弱学习器来构建一个强分类器。该算法会迭代调整训练样本的权重，更加关注被先前学习器错误分类的样本。每个弱学习器都在数据集的加权版本上进行训练，最终模型对这些弱学习器估计的

跑步模式类别进行加权投票。我们使用决策树作为弱学习器，其超参数优化方式与决策树相同，但估计器数量除外。

随机森林（RF）Breiman [2001], James 等人 [2013] 是一种集成机器学习方法，通过结合多个决策树的输出，克服其容易过拟合的倾向。随机森林在训练过程中使用数据和特征的随机子集，创建一组相互独立的决策树。随机森林的输出是由多数树选出的模式标签。我们对估计器数量（100至300）、最大深度（10至20）、每次分裂的最小样本数（2至10）以及叶节点的最小样本数（1至4）进行优化。我们评估了三种划分标准：基尼不纯度，以及基于香农信息论的两种度量——熵和逻辑损失。
人工神经网络（ANN）Bishop等人 [1995], James 等人 [2013] 是受人脑中生物神经网络的结构和功能启发的机器学习模型。它们由相互连接、按层组织的人工神经元组成。人工神经网络通过将输入变量传递至多个神经元层来预测跑步模式类别。每个神经元使用加权和与激活函数处理输入，后者为模型引入非线性。我们使用 ReLU 激活函数和 1 个网络层，并对神经元数量（1至100）和最大迭代次数（100至1000）进行优化。

卷积神经网络（CNN）LeCun等人 [2015], 是一种深度神经网络（DNN），通过一系列卷积层从输入数据中提取内在特征，即深度特征，使CNN在分类任务中表现出色。

时间序列的CNN架构如下：输入由两个相互独立的输入分支组成。a）主分支由形状为（T，A，C）的关节时序数据构成，其中 T 为时间步数量，A 为关节数量，C 为通道数量。b）辅助分支使用并行处理的稀疏特征。经过神经网络的输入a）后，下一层为标准差为0。05的高斯噪声层，用于正则化。完成正则化后，我们使用两个连续的基于 Inception 的残差块Szegedy等人 [2017] 进行特征提取，每个残差块包含64个滤波器。这些特征提取模块可实现多尺度处理，并通过残差连接保留信息。完成特征提取后，我们使用Squeeze-and-Excitation（SE）Hu 等人 [2018] 模块，突出最有用的特征通道，同时抑制无关响应。在第一个残差块之后，我们使用丢弃率为0。3的丢弃层，以减少过拟合。经过所有这些层后，将所得特征图展平，以获得输入时间序列的定长表示。

从次级分支开始，其中包含稀疏特征，这些特征相对于网络的第一部分进行并行处理。第二个分支由一个包含 8 个神经元并采用 L2 正则化的全连接层组成，随后依次进行批归一化、SiLU 激活和丢弃，丢弃率为 0。3。然后应用第二个包含 16 个神经元的全连接层，随后进行批归一化和 SiLU 激活。接着，将这些次级输出与从输入时间序列的第一种表示中获得的展平映射进行拼接。
最后，融合特征经过一个包含 16 个神经元并采用 L2 正则化的全连接层，随后进行批归一化和 SiLU 激活。网络的最后一层由一个带有 sigmoid 激活的全连接层组成，用于执行二分类。

训练使用 RMSprop 优化器，学习率为 10^{-4}，$\rho = 0.9$，并在 2 个步骤上进行梯度累积。我们采用二元交叉熵作为损失函数，并在训练过程中使用准确率、精确率和召回率作为评估指标。此外，我们在训练中使用了能够减轻类别不平衡效应的数据生成器。
长短期记忆（LSTM）由 Hochreiter 和 Schmidhuber 提出，是一种专门设计的循环神经网络（RNN）架构，旨在对序列数据中的长程依赖进行建模，同时缓解传统 RNN 固有的梯度消失问题。它通过记忆单元在时间推移中保持状态，并由三种门控机制加以调节：输入门控制流入单元的新信息；遗忘门决定保留或丢弃先前细胞状态的程度；输出门则控制内部状态向下一层或时间步的传递。这些门通过 Sigmoid 激活和逐元素乘法实现，从而能够精细控制信息的传播与保留。这种结构使长短期记忆网络能够有效学习复杂序列中的时间模式，例如金融时间序列、自然语言或生理信号。

在本研究中，我们拟使用按每个关节组织为时间序列的多变量时序数据。网络输入由一个张量构成，其长度固定为101个时间步，包含10个关节和9个通道，分别对应三种不同后处理模态各自的滚转、俯仰和偏航坐标。

本文提出的循环神经网络基于 LSTM 循环结构，首先采用一个包含 30 个滤波器、卷积核大小为 3 的双向一维卷积 LSTM 层，从而能够联合建模各关节之间的空间相关性以及跨时间的时间动态。我们返回处理后的序列，以保留后续处理所需的时间分辨率，并应用系数为 0.4 的循环丢弃。一旦

通过该层，我们将每个时间步的空间维度展平为单个特征向量。这些展平特征用于输入另一个包含 20 个神经元的双向LSTM层，其丢弃率为 0.2。在此情况下，我们不返回序列，而是获得一个表示潜在空间及其关联输入时序序列的单一输出。最后，潜在表示经过一个包含 10 个神经元并采用 ReLU 激活的全连接层处理，随后连接一个具有单个神经元并采用 Sigmoid 激活的稠密输出层。
训练使用 RMSprop 优化器，学习率为 10^{-4}、$\rho = 0.9$，并进行 2 步梯度累积。我们使用二元交叉熵作为损失函数，并在训练过程中以准确率、精确率和召回率作为评估指标。此外，我们在训练中使用数据生成器来缓解类别不平衡效应。

3.6 损伤风险检测的分类指标

我们使用四项分类指标评估模型在跑步模式分类中的表现：准确率（ACC）、精确率（PRE）、召回率（REC）和F1分数（$F1$）：

$$ACC(\%) = \frac{TP + TN}{P + N} \cdot 100 \tag{1}$$

$$PRE(\%) = \frac{TP}{TP + FP} \cdot 100 \tag{2}$$

$$REC(\%) = \frac{TP}{TP + FN} \cdot 100 \tag{3}$$

$$F1(\%) = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \cdot 100 \tag{4}$$

其中，TP、TN、FP、FN、P和 N分别表示真正例、真负例、假正例、假负例、正标签和负标签。本文中，负标签始终对应健康跑步模式类别，正标签则指损伤跑步模式类别，即 PFPS+ITBS、PFPS 或 ITBS，具体取决于分类问题。

我们使用五折交叉验证对所分析的分类器进行评估，并以志愿者为单位划分数据，以确保测试集中的志愿者未被纳入训练集。每轮迭代中，我们使用 80% 志愿者的时间序列和点值训练分类器，并使用其余 20% 志愿者的数据进行测试。所报告的分类指标为所有测试集指标的平均值及其标准差。

3.7 可解释性

在本研究中，我们采用多种方法来提高模型的可解释性，帮助临床医生更好地理解不同类型的运动与损伤之间的关系。这些方法包括基于特征的方法（例如 Shapley值），以及基于模型的技术，例如显著性图和 Grad-CAM。

Shapley值 Štrumbelj 和 Kononenko [2013] 用于量化每个特征对模型预测的贡献，方法是考察特征的所有可能组合。从数学角度看，某个特征的 Shapley值定义为该特征在所有可能的特征子集中的平均边际贡献。这种方法基于合作博弈论，其中每个特征都被视为一场博弈中的“参与者”，而“收益”则是模型的输出。由此得到的特征重要性得分分配公平且一致，反映了每个特征在不同情境下对预测的贡献程度。

显著性图通过识别并可视化对模型决策影响显著的输入区域，增强了深度神经网络等复杂系统的可解释性。生成显著性图的方法是计算输出相对于每个输入特征的导数，从而突出显示对预测影响最大的区域 Simonyan 等人 [2013]。这样，与某个类别相关的输出预测中受影响更大的区域就会被突出显示。

Grad-CAM Selvaraju 等人 [2017]是一种可视化技术，可帮助我们了解输入信号的哪些部分对模型的最终决策影响最大。这种方法主要为卷积神经网络设计，其原理是对输出相对于卷积层激活图的梯度进行加权组合，该卷积层通常位于 Softmax 层之前。输出结果是一张热力图，用于突出显示对模型预测贡献最大的输入区域。
显著性图和 Grad-CAM 对于其可视化结果符合人类直觉的输入非常有用，这类输入通常是图像，但它们也可迁移到其他情况，例如本文所讨论的情况：将所使用的信息表示为二维图，为检测损伤模式提供线索。我们使用热力图表示每个角度时间序列的累积影响，以分析它们对检测髌股疼痛综合征和髂胫束综合征的影响。

表 1：根据分类器中的输入参数，统计 PFPS+ITBS、PFPS 和 ITBS 的检测准确率。准确率以百分比表示。每组输入和损伤模式检测中，准确率最高的两项指标以粗体标出。

模型 髌股疼痛综合征+髂胫束综合征	时间序列			时间序列和点值			点值		
		PFPS	ITBS	髌股疼痛综合征+髂胫束综合征	PFPS	ITBS	髌股疼痛综合征+髂胫束综合征	PFPS	ITBS
KNN	55.8 ± 2.0	62.4 ± 3.0	60.0 ± 6.5	57.6 ± 4.5	64.6 ± 3.3	59.4 ± 4.2	$\mathbf{61.2 \pm 4.8}$	67.3 ± 5.4	62.5 ± 6.8
SVM_L	$\mathbf{62.2 \pm 4.8}$	$\mathbf{71.5 \pm 5.3}$	61.1 ± 7.6	$\mathbf{63.4 \pm 3.4}$	$\mathbf{71.7 \pm 4.2}$	$\mathbf{60.5 \pm 9.5}$	58.0 ± 7.4	67.4 ± 4.3	56.8 ± 7.5
SVM_P	58.3 ± 8.2	64.9 ± 4.6	55.0 ± 3.1	60.3 ± 7.9	63.7 ± 3.0	54.1 ± 4.2	56.7 ± 8.0	62.8 ± 4.7	56.1 ± 6.1
GP	$\mathbf{63.7 \pm 5.5}$	61.1 ± 10.2	$\mathbf{63.5 \pm 10.8}$	$\mathbf{66.0 \pm 3.8}$	65.3 ± 8.0	60.1 ± 7.1	$\mathbf{62.8 \pm 6.3}$	$\mathbf{69.5 \pm 3.7}$	$\mathbf{62.5 \pm 5.3}$
DT	53.7 ± 7.5	55.8 ± 8.7	53.5 ± 8.2	54.5 ± 5.3	51.6 ± 7.5	56.0 ± 7.5	49.7 ± 2.8	56.4 ± 4.3	51.0 ± 5.3
ADB	60.9 ± 6.1	$\mathbf{67.6 \pm 7.1}$	56.5 ± 6.9	58.8 ± 5.2	64.9 ± 5.9	58.9 ± 13.9	60.6 ± 8.0	62.3 ± 6.3	61.6 ± 8.3
RF	61.5 ± 3.2	65.3 ± 2.1	$\mathbf{63.6 \pm 8.3}$	60.4 ± 9.1	67.2 ± 4.7	$\mathbf{61.4 \pm 7.3}$	60.0 ± 6.5	65.5 ± 4.7	61.7 ± 5.8
ANN	58.0 ± 2.9	67.5 ± 4.0	61.1 ± 10.7	58.4 ± 4.4	$\mathbf{72.1 \pm 7.2}$	59.6 ± 8.4	60.5 ± 6.9	$\mathbf{69.7 \pm 7.0}$	$\mathbf{65.3 \pm 10.2}$
CNN	$\mathbf{71.4 \pm 1.8}$	$\mathbf{76.2 \pm 1.8}$	$\mathbf{73.8 \pm 2.0}$	$\mathbf{70.5 \pm 1.7}$	$\mathbf{77.9 \pm 1.2}$	$\mathbf{73.3 \pm 1.6}$	–	–	–
LSTM	65.1 ± 1.5	71.5 ± 1.6	66.7 ± 2.5	–	–	–	–	–	–

在我们的方法中，这些可解释性技术应用于性能最佳的机器学习和深度学习模型。具体而言，我们将结果表示为测试集某一折中阳性病例的平均响应。显著性图和 Grad-CAM 虽然最初是为图像数据设计的，但我们通过将输入时间序列表示为二维图，将其适配到我们的研究情境中。这样，我们就能突出显示对模型预测影响最大的时间区域和关节角度。由此得到的热力图为理解与髌股疼痛综合征和髂胫束综合征相关的生物力学模式提供了有价值的洞见，并为临床医生提供可解释的视觉线索，有助于支持损伤诊断和预防策略。

4 结果与讨论

4.1 优化的超参数

通过网格搜索确定了各分类模型的最优超参数。对于K近邻算法，采用欧氏距离度量和 $k = 7$时性能最佳。SVM_L 在正则化参数为0.001时结果最优，而SVM_P 在正则化参数为1.1、多项式次数为3时表现最佳。决策树模型采用随机划分策略进行优化，最大深度为10，最大特征数设为平方根，每个叶节点的最小样本数为2，每次分裂的最小样本数为2。对于随机森林分类器，最佳配置包括200棵决策树、最大深度为10、每次分裂的最小样本数为2、每个叶节点的最小样本数为1、最大特征数采用以2为底的对数、以基尼熵作为分裂准则，并使用自助采样。最后，人工神经网络在包含100个神经元的单个隐藏层、最大训练迭代次数为1000时达到最佳性能。

4.2 模型表现

模型检测与损伤相关的跑步模式时，其准确率取决于输入集和具体损伤。表 1 展示了所有方法中经评估模型的准确率。在所有条件下，深度学习算法都取得了最高准确率。其中，卷积神经网络表现最佳：使用时间序列和点值时，对髌股疼痛综合征的准确率达到 77.9%；使用时间序列时，对髂胫束综合征的准确率达到 73.8%；在相同输入类型下，对综合损伤类别（髌股疼痛综合征+髂胫束综合征）的准确率达到 71.4%。在机器学习方法中，支持向量机$_L$ 和高斯过程在大多数条件下准确率最高，因此是最具竞争力的经典模型。在损伤模式识别中，高斯过程在所有输入数据组合下都表现稳定；使用全部输入数据时，对髌股疼痛综合征+髂胫束综合征的准确率最高（使用时间序列时为 63.7%，使用时间序列和点值时为 66.1%，使用点值时为 62.5%）。使用点值时，高斯过程尤为有效：对髌股疼痛综合征+髂胫束综合征的准确率为 62.8%，对髌股疼痛综合征为 69.5%，对髂胫束综合征为 62.5%。同样，支持向量机$_L$ 表现稳健，在时间序列与点值组合下取得了一些最高准确率：检测髌股疼痛综合征+髂胫束综合征的准确率为 63.4%，对髌股疼痛综合征为 71.7%，对髂胫束综合征为 60.5%。鉴于这些表现，以下讨论将重点介绍高斯过程和支持向量机$_L$，因为它们在机器学习方法中表现优异；同时也将讨论作为深度学习领先方法的卷积神经网络。

输入数据是否充分，很大程度上取决于算法，即使仅考虑 GP、SVM_L 和卷积神经网络。对于髌股疼痛综合征检测，结合时间序列与点值后，准确率略有提升：SVM_L 的准确率从 71.5% 升至 71.7%，卷积神经网络则从 76.2% 升至 77.9%。对于髌股疼痛综合征+髂胫束综合征，GP 的准确率从 63.7% 升至 66.0%，SVM_L 也呈现类似趋势（从 62.2% 升至 63.4%），而卷积神经网络的准确率略有下降（从 71.4% 降至 70.5%）。相比之下，在髂胫束综合征模式识别中，单独使用点值或时间序列效果最佳。GP 使用时间序列时的准确率为 63.5%，使用点值时为 62.5%，而两者结合后准确率降至 60.1%。使用卷积神经网络时，髂胫束综合征的最高准确率（73.8%）也来自时间序列。

表 2：精确率、召回率 以及损伤模式识别中最准确方法的 F1分数分类指标 .

		时间序列			时间序列与点值			点值		
	模型	PFPS+ITBS	PFPS	ITBS	PFPS+ITBS	PFPS	ITBS	PFPS+ITBS	PFPS	ITBS
$PRE(\%)$	SVM_L	61.9 ± 3.9	$\mathbf{69.9 \pm 6.3}$	$\mathbf{62.4 \pm 7.8}$	63.2 ± 2.2	$\mathbf{70.4 \pm 2.1}$	$\mathbf{60.3 \pm 6.2}$	57.3 ± 5.3	66.8 ± 4.9	56.5 ± 8.9
	GP	$\mathbf{64.4 \pm 4.9}$	61.3 ± 14.8	$\mathbf{62.3 \pm 9.2}$	$\mathbf{66.5 \pm 2.3}$	64.6 ± 9.3	59.3 ± 3.4	$\mathbf{62.8 \pm 4.0}$	$\mathbf{68.4 \pm 5.1}$	$\mathbf{63.8 \pm 5.3}$
	CNN	$\mathbf{52.9 \pm 3.5}$	$\mathbf{42.2 \pm 2.5}$	$\mathbf{40.4 \pm 4.9}$	52.2 ± 1.4	41.3 ± 3.0	34.3 ± 3.0	–	–	–
	LSTM	43.8 ± 1.0	37.4 ± 2.2	26.8 ± 2.7	–	–	–	–	–	–
$REC(\%)$	SVM_L	$\mathbf{66.5 \pm 7.2}$	76.4 ± 9.5	55.3 ± 16.8	$\mathbf{66.5 \pm 6.9}$	77.0 ± 2.5	$\mathbf{61.3 \pm 18.8}$	59.3 ± 16.5	69.1 ± 8.9	50.2 ± 27.7
	GP	63.1 ± 3.7	$\mathbf{80.5 \pm 16.2}$	$\mathbf{60.7 \pm 19.9}$	65.9 ± 2.4	$\mathbf{78.6 \pm 11.4}$	59.9 ± 17.1	$\mathbf{63.0 \pm 8.1}$	$\mathbf{73.1 \pm 5.4}$	$\mathbf{60.0 \pm 6.9}$
	CNN	82.2 ± 6.6	$\mathbf{82.7 \pm 8.1}$	72.5 ± 10.2	79.7 ± 7.3	90.3 ± 5.8	81.87 ± 4.6	–	–	–
	LSTM	$\mathbf{94.4 \pm 3.6}$	77.8 ± 8.1	$\mathbf{87.5 \pm 6.6}$	–	–	–	–	–	–
$F1(\%)$	SVM_L	$\mathbf{63.8 \pm 3.6}$	$\mathbf{72.4 \pm 4.7}$	57.2 ± 10.8	64.6 ± 3.0	$\mathbf{73.5 \pm 1.4}$	$\mathbf{59.4 \pm 11.3}$	57.4 ± 9.8	67.8 ± 6.3	49.0 ± 16.1
	GP	63.7 ± 4.0	66.6 ± 6.0	$\mathbf{60.7 \pm 14.7}$	$\mathbf{66.2 \pm 2.1}$	69.5 ± 3.9	58.6 ± 10.1	$\mathbf{62.8 \pm 5.9}$	$\mathbf{70.5 \pm 4.1}$	$\mathbf{61.4 \pm 3.9}$
	CNN	$\mathbf{64.1 \pm 1.7}$	$\mathbf{55.7 \pm 1.4}$	$\mathbf{51.2 \pm 3.0}$	62.9 ± 2.1	56.5 ± 2.1	48.2 ± 1.6	–	–	–
	LSTM	59.8 ± 1.2	50.3 ± 1.7	40.8 ± 2.6	–	–	–	–	–	–

表 2 展示了准确率最高的模型——GP、SVM_L 和卷积神经网络——的精确率、召回率和 F1分数指标，以便进行更详细的分析。使用点值时，GP 表现突出，在所有指标上均优于 SVM_L；对于髌股疼痛综合征+髂胫束综合征或髌股疼痛综合征，各项指标提升幅度为 2% 至 5%，而髂胫束综合征的 F1分数最高提升 11%。然而，在处理时间序列或结合时间序列与点值时，SVM_L 的精确率和 F1分数更高，而 GP 的召回率仍占优势。尽管两种模型的精确率、召回率和 F1分数相近，GP 始终较高的召回率值得注意，因为这可能有助于更好地识别真实的损伤模式。对于卷积神经网络，还可以观察到其他特点：由于优先考虑召回率，其精确率相对较低。这意味着，尽管总体准确率较高，其 F1分数仍有提升空间。

这些结果表明，与未受伤模式相比，深度神经网络能够捕捉损伤跑步模式的动作捕捉时间序列数据中的复杂差异，但其优化目标是最大化召回率，精确率仍有提升空间。与点值相比，分析时间序列数据时结果有所改善，这支持了以下观点：时间序列包含与识别跑步者受伤相关模式有关的运动信息 Phinyomark 等人 [2017]。支持向量机$_L$也能有效处理时间序列和点数据，以识别损伤跑步模式，而高斯过程更适合使用点值。

与我们此前在 Fuentes-Jiménez 等人 [2025], 中的研究相比，报告的指标总体有所下降。尽管两项研究使用相同的数据库，但本研究采用了更严格的数据划分策略：测试子集和训练子集中的用户完全分离。与先前的方法不同，在本研究评估模型时，测试数据不包含任何来自训练数据中志愿者的样本。这一区别使当前分析更加严谨，所得结果也更能反映一种现实情境：分类器被用于检测其从未接触过的跑步者的损伤模式。

4.3 损伤模式检测

与髌股疼痛综合征相关的损伤模式检测准确率最高。结合时间序列和点值时，支持向量机$_L$和人工神经网络在髌股疼痛综合征模式识别中的准确率均超过 70%（见表 1）；卷积神经网络的表现更佳，使用相同输入时准确率达到 77.9%。除仅使用点值的情况外，髌股疼痛综合征+髂胫束综合征通常是检测效果第二好的模式。对于髌股疼痛综合征+髂胫束综合征，在使用时间序列时，大多数机器学习方法的检测准确率介于 60% 至 66% 之间；而卷积神经网络使用时间序列时，髂胫束综合征的检测准确率提升至 73.8%，结合时间序列和点值时则达到 73.3%。这些结果支持了以下假设：分析完整时间序列，而非将其简化为点运动值，与识别损伤模式息息相关。

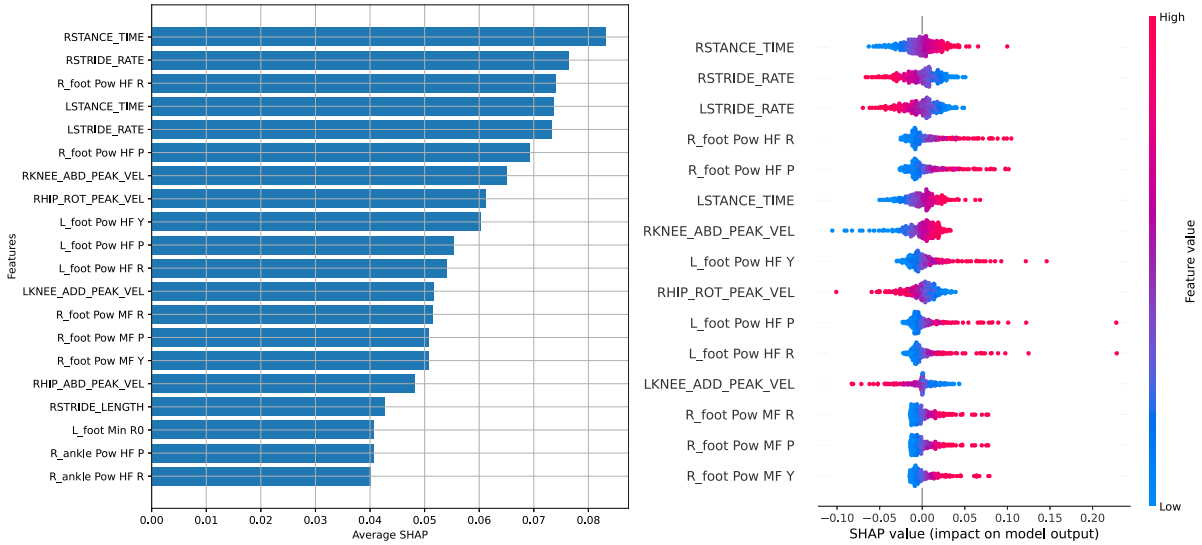
分析 F1分数时也得出了相同结论；F1分数体现了精确率与召回率之间的平衡。在机器学习算法中，髌股疼痛综合征的 F1分数最高，其次是一般情况（髌股疼痛综合征+髂胫束综合征），最后是髂胫束综合征（见表 2）。然而，卷积神经网络呈现出不同的模式：其 F1 分数在髌股疼痛综合征 + 髂胫束综合征这一情形下最高，因为对于髌股疼痛综合征和髂胫束综合征特异性损伤，为了尽可能检测出损伤模式，精确率会显著降低。在检测髌股疼痛综合征时，准确率达到77.9%，这表明卷积神经网络识别出了受伤者与未受伤志愿者在角度时间序列和运动参数上的差异。对于髂胫束综合征，识别也利用了某些差异，但准确率（73.8%）仍有较大的提升空间。与此前无法使用无监督算法对受伤和未受伤志愿者进行分组的研究相比，这些结果具有创新性（Dingenen 等人 [2020], Jauhiainen 等人 [2020], Senevirathna 等人 [2023]）。因此，本研究结果表明，先进算法能够利用运动分析系统识别跑步中的损伤模式。

与将髌股疼痛综合征模式与其他模式结合起来相比，单独考察时对该模式的检测更准确。这支持了这样的观点：跑步模式的差异与特定损伤的关联更为密切，而非与一般损伤模式相关，这与此前的研究结果一致（Ceyssens 等人 [2019]）。此外，与髂胫束综合征相比，髌股疼痛综合征模式的检测效果更好，这表明髌股疼痛综合征模式与未受伤模式的偏离程度最大。

本研究的分类指标与这一观点一致：机器学习方法能够成功识别损伤风险较高的运动员。本研究中用于识别髌股疼痛综合征（PFPS）的指标低于 Eskofier 等人[2012], 报告的结果，后者显示准确率达到 100%。不过，先前的研究仅识别出 6 名受伤志愿者，并确保未受伤组的特征相似。相比之下，我们的研究纳入了更多变异性，这可以解释分类指标较低的原因。先前研究报告称，在预测棒球或澳式足球等其他运动项目的损伤风险时，分类准确率同样较高（75% 至 83%）（Hans Van Eetvelde 等人[2021], ）；临床步态分析领域也报告了类似结果，并提倡运用机器学习辅助医疗决策（Slijepcevic 等人[2021]）。

4.4 可解释性

我们使用累积沙普利值评估机器学习算法的可解释性，以识别最重要的输入特征并分析其 SHAP 贡献。我们重点分析支持向量机（SVM）$_{\text{L}}$的特征重要性，因为该方法以时间序列和点值作为输入数据时均能持续取得较高性能，从而使我们能够在最佳条件下评估所有参数的重要性。我们针对每种损伤模式检测分别进行分析，结果见图 2a、图 3a 和图 4a，分别对应 PFPS+IBTS、PFPS 和 ITBS。

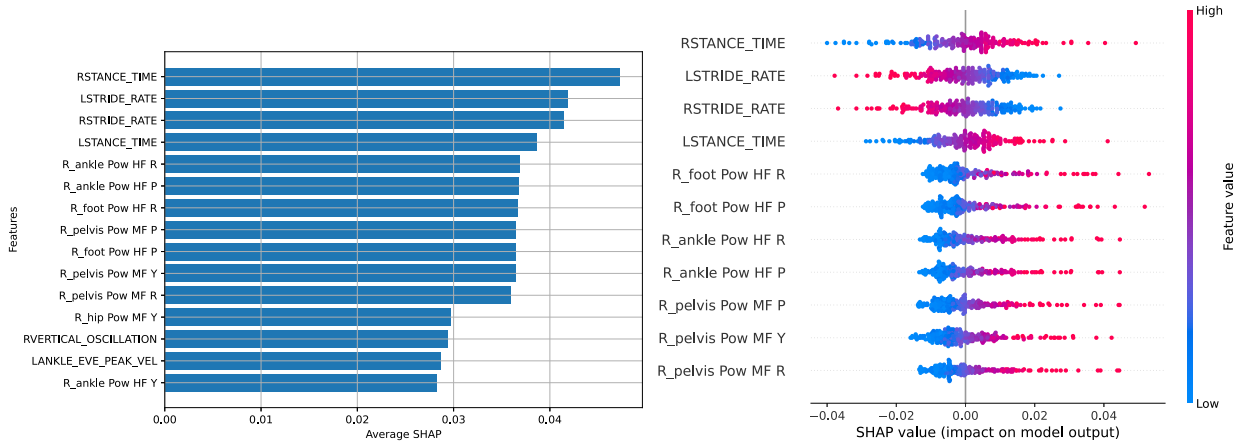


（a）累积 SHAP 值。　　（b）最重要特征的 SHAP 值。

图 2：累积归一化　用于使用支持向量机识别 PFPS+ITBS 的最相关特征的 SHAP 值$_{\text{L}}$.

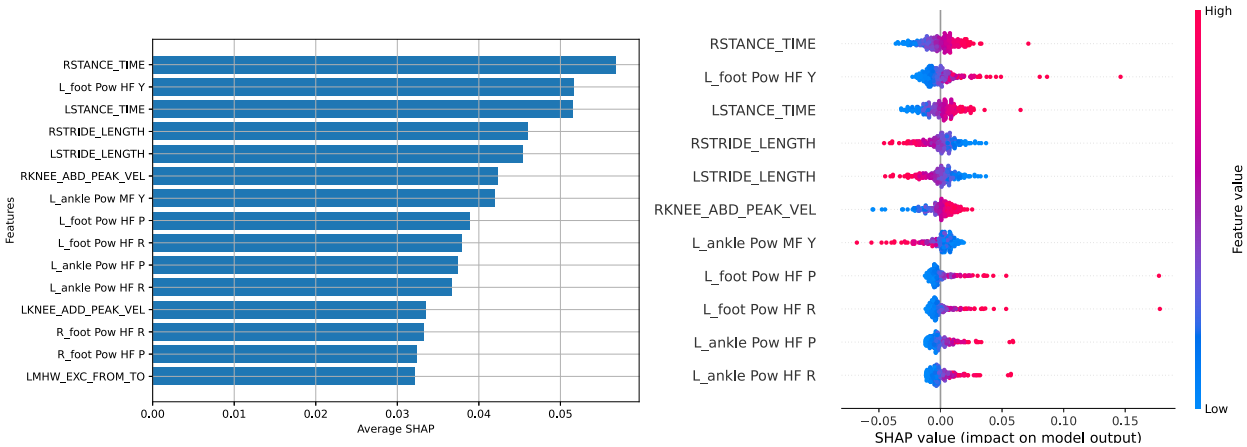
在目视检查图 2a、图 3a 和图 4a 后，我们将边界设定为研究第二次 SHAP 值递减中的最相关特征。分析结果表明，对于 PFPS+ITBS 检测，我们分析 SHAP 值大于或等于 0.05 的特征；对于 PFPS，分析 SHAP 值大于或等于 0.03 的特征；对于 ITBS，分析 SHAP 值大于或等于 0.35 的特征。通过这种方式，我们在图 2b 中分析 PFPS+ITBS 的五个测试集所累积的 15 个特征的 SHAP 值、图 3b 中 PFPS 的 11 个特征的 SHAP 值，以及图 4b 中 ITBS 的 11 个特征的 SHAP 值。

分析 SHAP 值最高的特征后，我们发现，图 2b 所示的通用情形中的相关特征，通常兼具了分别针对 PFPS 和髂胫束综合征所识别的重要特征（见图 3 和图 4）。通用损伤模式和特定损伤模式中均包含的相关特征示例包括双脚的支撑时间（对 PFPS 和髂胫束综合征而言分别也具有显著性）、步频（与 PFPS 相关）以及膝外展峰值速度（与髂胫束综合征相关）。点值（包括中频和高频频带中的功率）表现出更高的重要性；尽管在纳入时间序列数据时，支持向量机$_{\text{L}}$ 取得了更好的结果，这表明这些时间特征虽具有相关性，但与点值结合时并非最具决定性的因素。这些特征的高值或低值的影响，以及其对模型影响的方向，在通用情形和个体情形中保持一致。例如，在所有三种情形（PFPS+髂胫束综合征、



（a）累积 SHAP 值。 （b）最重要特征的 SHAP 值。

图 3：最相关特征的累积 归一化 SHAP 值，用于通过支持向量机识别髌股疼痛综合征$_L$.



(a) SHAP值累积。 (b) 最重要特征的 SHAP值。

图 4：使用支持向量机$_L$检测髂胫束综合征的累积归一化 SHAP。

PFPS 和髂胫束综合征识别）中，双脚较长的支撑时间都与损伤风险直接相关。这一发现表明，PFPS 和髂胫束综合征的模式有所不同；然而，机器学习方法能够分析通用损伤模式，并检测出与损伤关联最强的参数。

髌股疼痛综合征和髂胫束综合征的损伤模式都以支撑时间为共同特征；然而，其余特征及由此形成的整体模式有所不同。髌股疼痛综合征的损伤模式与较长的支撑时间和较低的步频相关。还可以观察到，足部、踝关节和骨盆在中、高频带上的高功率与这种损伤风险有关。髂胫束综合征与支撑时间仍保持相同的直接关系，但步长也有影响；步长越长，与髂胫束综合征损伤的关联越弱。运动的高频和中频成分也与这种损伤有关；不过在这种情况下，最重要的关节仅为足部和踝关节，而不像髌股疼痛综合征那样还包括骨盆。另一项差异是膝外展峰值速度的重要性：较低的数值与髂胫束综合征风险降低的关联更强。

采用时间序列的卷积神经网络的可解释性图见图5（一般情况）、图6（髌股疼痛综合征）和图7（髂胫束综合征）。对于所有损伤，我们都展示了显著性图、Grad-CAM和SHAP值，以分析它们的一致性。在检测三种损伤模式——髌股疼痛综合征+髂胫束综合征、髌股疼痛综合征和髂胫束综合征——时，所评估的三种图之间具有一致性：具有显著性的区域也呈现较高的Smooth-Grad值，以及接近1且幅度较大的正SHAP值。例如，在图5所示的一般情况下，步态周期中20%至40%的时间区间，在身体左侧的关节上始终呈红色，这意味着这些关节与损伤模式的识别有关。相反，对损伤检测相关性较低的区间，例如身体左侧的初始阶段，

在三种图中均呈蓝色，表明显著性较低（图5a）、Smooth-Grad值较低（图5b），且SHAP值为负（图5c）。

一般情况与 PFPS 和 ITBS 的具体情况之间存在显著差异。一般情况下（其可解释性见图 5），相关关节数量最多且持续时间最长的区间出现在支撑中期，从支撑期的 20% 持续至 60%。在此区间内，步态这一阶段的开始或结束时，双足都显示为相关部位。在步态开始和结束时也检测到相关性，但持续时间较短。PFPS（图 6）中可以观察到两个明显的相关区域，尤其是从 SHAP值来看：支撑期开始阶段（至 20%）以及接近支撑期结束的时点（支撑中期结束）。在步态开始时，左踝关节显示出较高相关性，双足的相关性则更为显著，这凸显了初始接触时足部放置的重要性。在支撑中期结束时，最相关的关节是腿部上方的关节，具体而言是双腿的膝关节和髋关节。最后，在 ITBS 中，图 7 显示相关性最高的部位是支撑期开始阶段，此时身体两侧的足部和踝关节都很重要，与 PFPS 类似。在支撑中期，相关性转向位置更高的关节，例如髋关节，这与 PFPS检测的情况类似。

在我们的分析中，这些可解释性图无法确定与 PFPS 或髂胫束综合征相关的明确、特定的跑步模式。不过，显著性图、Grad-CAM 和 Shapley图有助于识别最相关的时间区间。总结本分析中特有的模式，通用情形（图 5）涵盖整个支撑期，同时改变相关关节或身体节段，从而覆盖所有关节或节段。相比之下，与 PFPS（图 6）和髂胫束综合征（图 7）最相关的区间，显然是支撑期开始时小腿下部的运动，以及支撑中期结束时的运动。这些信息对未来致力于检测此类损伤模式的研究具有价值。此外，通用情形、PFPS 和髂胫束综合征的图在这些相关区间以及对分析影响最大的关节方面存在差异，这清楚地表明，所评估的两种损伤对应的跑步模式并不相同。这些发现为建立通用和特定的跑步损伤模式提供了令人鼓舞的结果。

5 结论

本研究利用监督式机器学习框架（包括经典方法和深度神经网络方法），评估了基于动作捕捉数据检测与髌股疼痛综合征和髂胫束综合征相关跑步模式的效果。卷积神经网络和长短期记忆网络在检测这些跑步模式方面优于经典机器学习模型。髌股疼痛综合征检测（77.9%）表明，卷积神经网络能够有效识别受伤跑步者与未受伤跑者在角度时间序列和运动参数上的差异。髂胫束综合征检测（73.8）也显示出一定差异，但差异程度低于髌股疼痛综合征检测。研究结果凸显了区分损伤模式的研究潜力。髌股疼痛综合征+髂胫束综合征识别分析（71.4%）也展现出良好结果，表明可以找到一种通用的损伤模式。此外，基于旋转角度时间序列训练的模型始终优于仅使用运动学描述符的模型，凸显了分析完整运动过程的技术的重要性。

研究使用 Shapley、显著性图和 Grad-CAM 评估分类算法的可解释性，以解读模型的决策。在对经典机器学习算法的分析中，髌股疼痛综合征和髂胫束综合征相关特征存在显著差异；在通用情境下，始终能找到这些特定情形特征的组合。分析卷积神经网络中完整时间序列的相关性时发现，髌股疼痛综合征与髂胫束综合征损伤模式对应的相关时间区间也各不相同，这证明动作捕捉系统能够测量这些模式之间的细微差异。

总之，本研究支持在跑步分析中使用可解释模型来分析运动学。这为未来开发个性化损伤预防工具提供了令人期待的结果。未来的研究可以探索其他深度神经网络技术，进一步挖掘特征之间的关系，并分析特征顺序及其填充的影响。此外，未来研究应尝试将此分析推广至便携式传感技术，从而能够在真实环境中更具代表性地评估跑步生物力学。

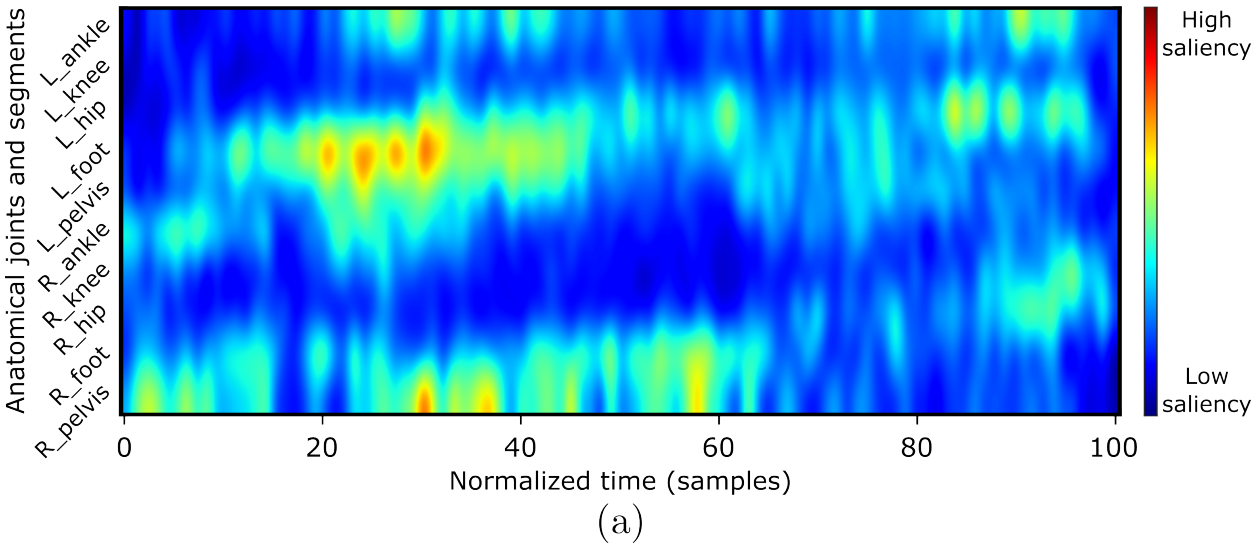
参考文献

S Videbaek、AM Bueno、RO Nielsen 和 S Rasmussen。不同类型跑步者每跑步 1000 小时的跑步相关损伤发生率：系统综述与荟萃分析。《运动医学》，45(7):1017–1026，2015年。

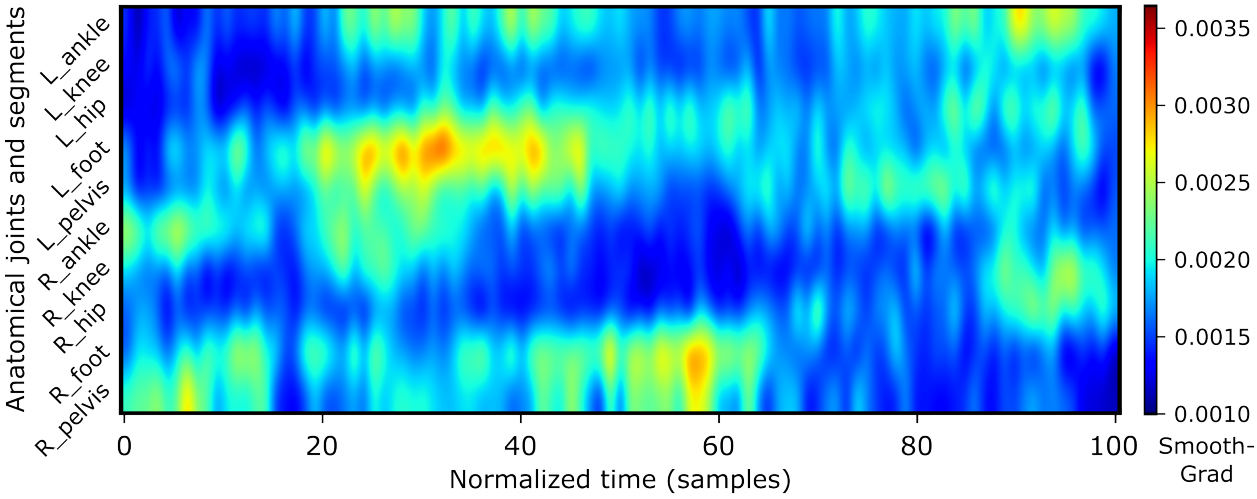
JE Taunton、MB Ryan、DB Clement、DC McKenzie、DR Lloyd-Smith 和 BD Zumbo。对 2002年跑步损伤的回顾性病例对照分析。《英国运动医学杂志》，36(2):95–101，2002年。

MP Van der Worp、DS ten Haaf、R van Cingel、A de Wijer、MW Nijhuis-van der Sanden 和 JB Staal。跑步者损伤：关于风险因素和性别差异的系统综述。《公共科学图书馆·综合》，10(2):e0114937，2015年。

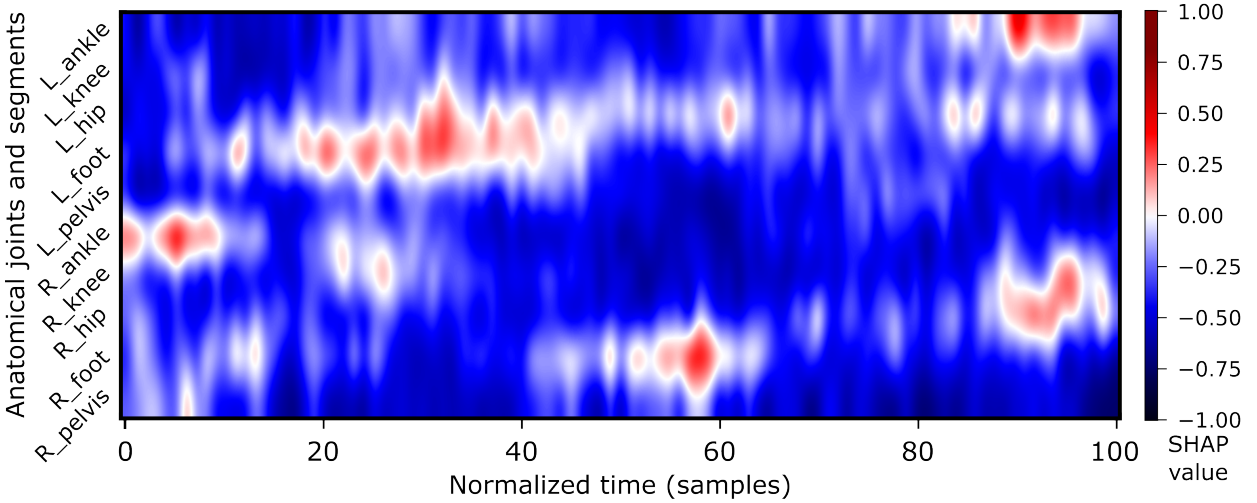
Harrison Philip Crowell 和 Irene S Davis。通过步态再训练降低跑步者的下肢负荷。《临床生物力学》，26(1):78–83，2011年。



(a)

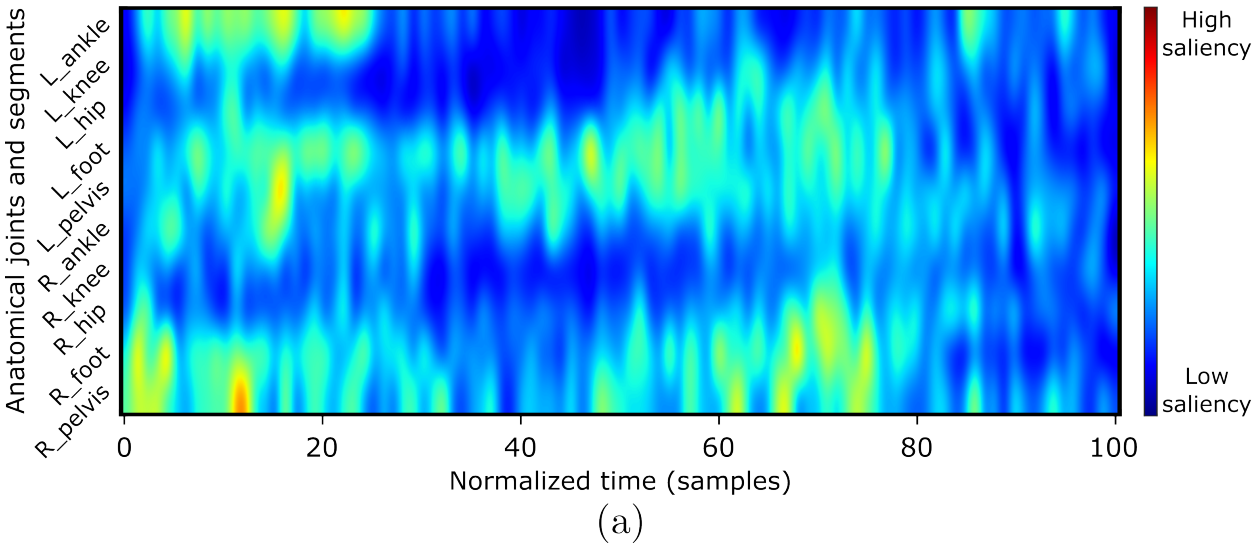


(b)

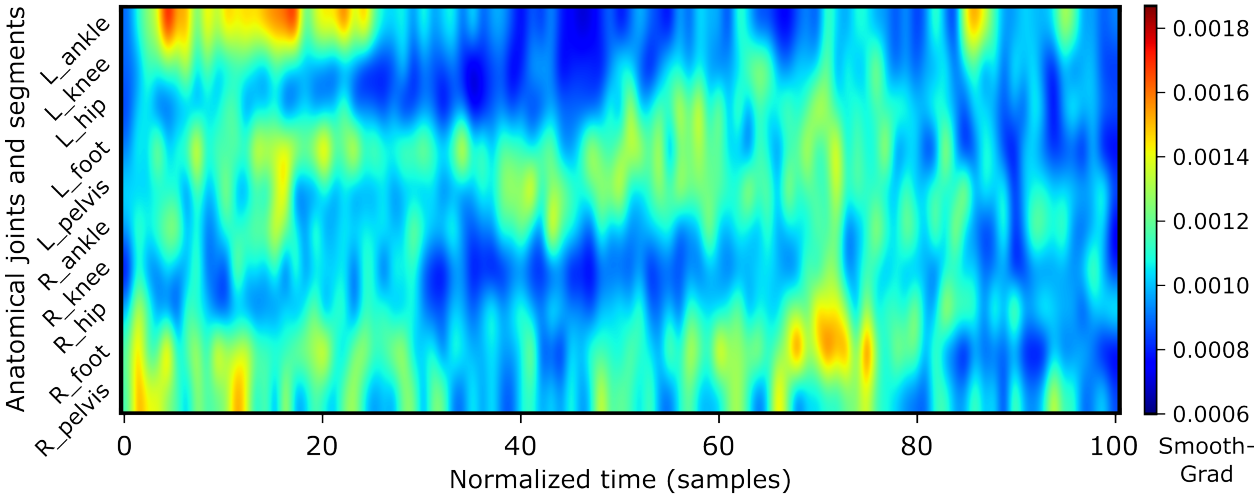


(c)

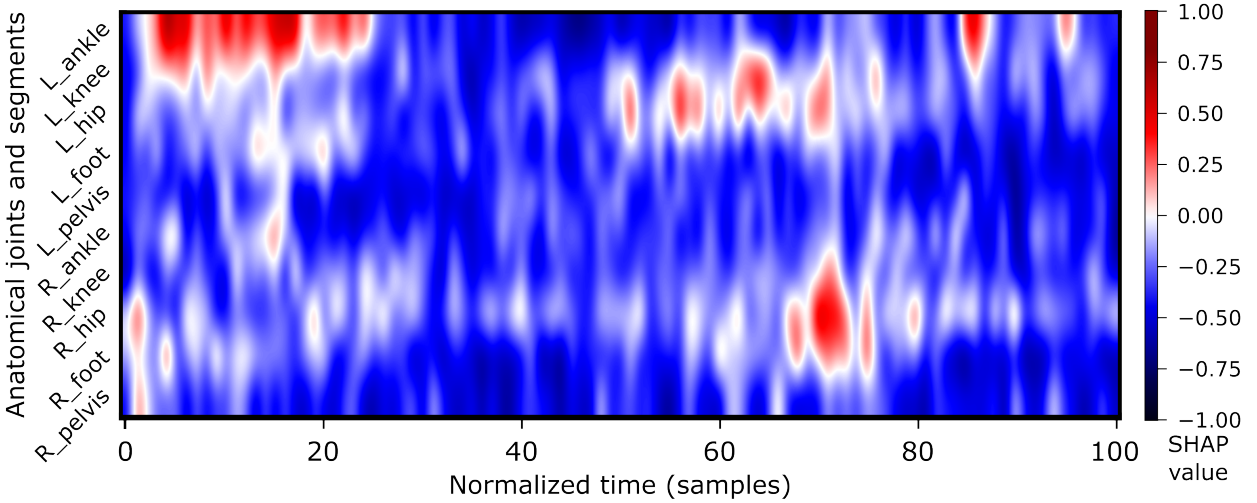
图 5：利用三种方法提取的时间特征进行 PFPS+ITBS 识别的卷积神经网络可解释性：(a) 为显著性图，(b) 对应 Grad-CAM，(c) 为 Shapley 分析结果。13



(a)

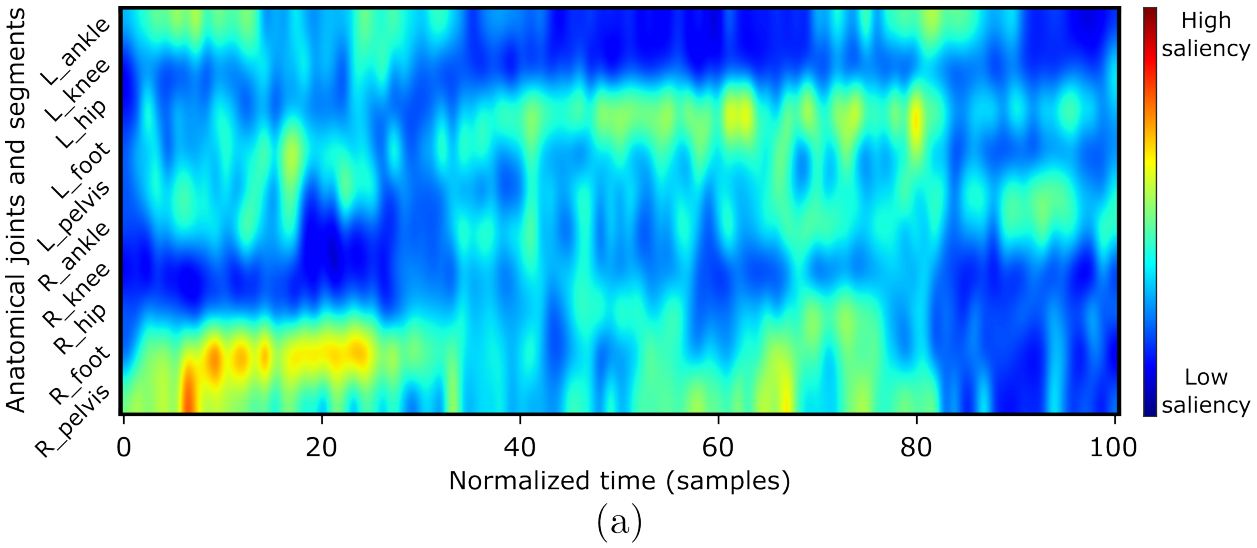


(b)

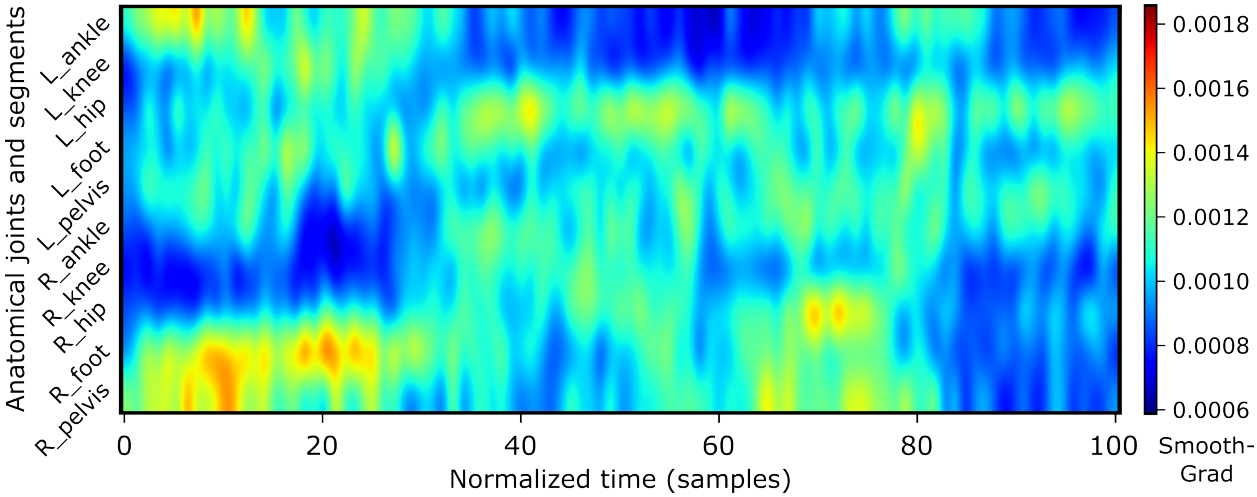


(c)

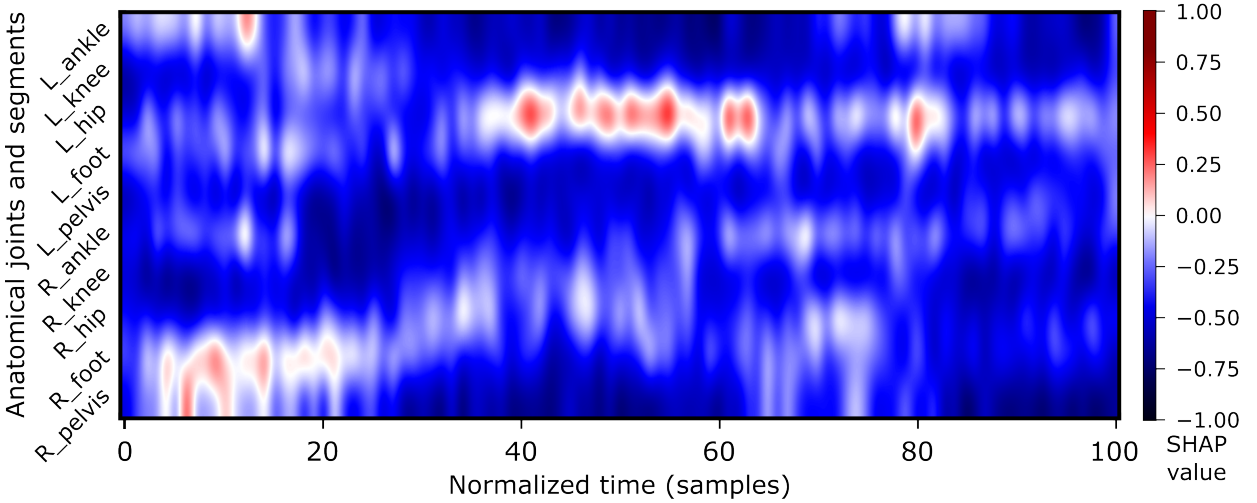
图 6：利用三种方法从时序特征中识别髌股疼痛综合征的卷积神经网络可解释性：(a) 显示显著性图，(b) 对应 Grad-CAM，(c) 为 Shapley 分析结果。14



(a)



(b)



(c)

图 7：使用三种方法提取的时序特征进行 ITBS 识别时，卷积神经网络的可解释性：(a) 显示显著性图，(b) 对应 Grad-CAM，(c) 为 Shapley 分析结果。15

Saghar Soltanabadi、Hooman Minoonejad、Mohammad Bayattork 和 Mohammad Seyedahmadi。虚拟现实与增强现实训练对预防运动员前交叉韧带损伤及加速康复的影响：范围综述。《亚洲运动医学杂志》，14(4)，2023年。Meghan Keast、Jason Bonacci 和 Aaron Fox。步态干预对跑步期间胫骨负荷的急性影响：系统综述与荟萃分析。《运动医学》，52(10):2483–2509，2022年。Grazia Cicirelli、Donato Impedovo、Vincenzo Dentamaro、Roberto Marani、Giuseppe Pirlo 和 Tiziana R D'Orazio。神经退行性疾病中的人体步态分析：综述。*IEEE* 生物医学与健康信息学杂志，26 (1):229–242，2021a。Mohammed Hammoud、Aleksei Shcherbak、Olga Istrakova、Nataliya Shindryaeva、Ekaterina Bril、Roberto Passerone 和 Andrey Somov。腕戴式传感器与机器学习在帕金森病检测中的应用：二分类与多分类问题研究。*IEEE*《仪器与测量汇刊》，2025年。Aleksei Shcherbak、Ekaterina Kovalenko 和 Andrey Somov。通过可穿戴传感器与机器学习检测和分类帕金森病早期阶段。*IEEE*《仪器与测量汇刊》，72:1–9，2023年。Hoonyong Lee、Gaang Lee、SangHyun Lee 和 Changbum R Ahn。基于置信区间估计预测的异常步态模式评估滑倒、绊倒和跌倒危险暴露。《建筑自动化》，139: 104253，2022年。Wei Zeng、Shiek Abdullah Ismail 和 Evangelos Pappas。基于步态动力学差异和神经网络检测前交叉韧带损伤。《人工智能评论》，53(5):3153–3176，2020年。Christos Kokkotis、Serafeim Moustakidis、Themistoklis Tsatalas、Charis Ntakolia、Georgios Chalatsis、Stylianos Konstadakos、Michael E Hantes、Giannis Giakas 和 Dimitrios Tsaopoulos。运用可解释机器学习识别与前交叉韧带损伤相关的步态生物力学参数。《科学报告》，12(1):6647，2022年。Yifang Zhuang、Wenxing Zhou、Ziwei Zeng、Shiwei Mo 和 Lin Wang。足部着地模式与地面条件对跑步期间下肢运动学和动力学的影响：统计参数映射分析。《公共科学图书馆·综合》，20(2):e0317853，2025年2月。ISSN 1932-6203。doi:10.1371/journal.pone.0317853。URL http://dx.doi.org/10.1371/journal.pone.0317853。Linde Ceyssens、Romy Vanelderen、Christian J Barton、Peter Malliaras 和 Bart Dingenen。与跑步相关损伤有关的生物力学风险因素：系统综述。《运动医学》，49(7):1095–1115，2019年。Angkoon Phinyomark、Giovanni Petri、Esther Ibáñez-Marcelo、Sean T. Osis 和 Reed Ferber。步态生物力学大数据分析：当前趋势与未来方向。《医学与生物工程学杂志》，38(2): 244–260，2017年7月。ISSN 2199-4757。doi:10.1007/s40846-017-0297-2。URL http://dx.doi.org/10.1007/s 40846-017-0297-2。Steffen Willwacher、Markus Kurz、Johanna Robbin、Matthias Thelen、Joseph Hamill、Luke Kelly 和 Patrick Mai。长跑运动员过度使用损伤的跑步相关生物力学风险因素：考虑损伤特异性和未来研究潜力的系统综述。《运动医学》，52(8):1863–1877，2022年3月。ISSN 1179-2035。doi:10.1007/s40279-022-01666-3。URL http://dx.doi.org/10.1007/s40279-022-01666-3。Bart Dingenen、Filip Staes、Romy Vanelderen、Linde Ceyssens、Peter Malliaras、Christian J Barton 和 Kevin Deschamps。基于标记点二维视频分析评估的跑步运动学，对有跑步相关损伤的休闲跑者进行亚分类。《运动物理治疗》，44:99–106，2020年。Susanne Jauhiainen、Andrew J Pohl、Sami Äyrämö、Jukka-Pekka Kauppi 和 Reed Ferber。通过层次聚类分析确定受伤跑步者是否呈现相似的运动学步态模式。《斯堪的纳维亚运动医学与科学杂志》，30:732–740，2020年。Jack A Martin、Stiffler-Joachim。Mikel R、Bryan C Wille 和 Christa Mand Heiderscheit。采用层次聚类方法考察跑步者骨应力性损伤的潜在风险因素。《生物力学杂志》，141:111136，2022年。Angela M Senevirathna、Andrew J Pohl、Matthew Jordan、William Brent Edwards 和 Reed Ferber。受伤与未受伤后足着地跑者之间的动力学变量差异：层次聚类分析。《斯堪的纳维亚运动医学与科学杂志》，33:160–168，2023年。Bjoern M Eskofier、Martin Kraus、Jay T Worobets、Darren J Stefanyshyn 和 Benno M Nigg。采用特征排序对运动学和动力学跑步数据进行模式分类，以区分性别、穿鞋／赤足状态及损伤组。《生物力学与生物医学工程中的计算方法》，15(5):467–474，2012年。

David Fuentes-Jiménez、Sara García-de Villa、David Casillas-Pérez、Pablo Floría 和 Francisco-Manuel Melgarejo- Meseguer。用于可解释地检测跑步者膝关节损伤的先进机器学习技术。载于*2025*年 *IEEE* 医学测量与应用会议 *(MeMeA)*，第1–6页，2025年。doi:10.1109/MeMeA65319.2025.11068035。Grazia Cicirelli、Donato Impedovo、Vincenzo Dentamaro、Roberto Marani、Giuseppe Pirlo 和 Tiziana R D'Orazio。神经退行性疾病中的人体步态分析：综述。*IEEE* 生物医学与健康信息学杂志，26 (1):229–242，2021b。Yi Xia 和 Aisheng Cao。用于检测步态冻结的多模态动静态组合图卷积网络。生物医学信号处理与控制，112:108368，2026年。M Ramkumar Raja、G Hari Krishnan、Amanulla Mohammad 和 Sreevardhan Cheerla。基于双时序门控多图卷积网络的神经退行性疾病检测与严重程度预测。生物医学信号处理与控制，112:108566，2026年。Yi Ru、Xiangwei Zhao、Yasamin Baghersad、Hamid Taheri Andani、Maboud Hekmatifar 和 Belgacem Bouallegue。步态分析与机器学习算法用于区分帕金森病、肌萎缩侧索硬化症和亨廷顿病。生物医学信号处理与控制，110:108216，2025年。Felipe Arruda Moura、Alexandre RM Pelegrinelli、Danilo S Catelli、Mario Lamontagne 和 Ricardo da Silva Torres。基于步态运动学预测膝骨关节炎：探索深度神经网络与迁移学习方法在时间序列分类中的潜力。《生物力学杂志》，第112881页，2025年。Sara García-de-Villa、Luisa Ruiz-Ruiz、Guillermo García-Villamil-Neira、Marta Álvarez Neira、Elisabet Huertas-Hoyas、Antonio J Del-Ama、M Crtistina Rodriguez-Sanchez、Fernando Seco 和 Antonio R Jiménez。验证一种基于惯性测量单元的步态分析方法，以传统方法为参照评估跌倒风险。*IEEE* 生物医学与健康信息学杂志，2024年。Yuchao Liu、Jiajie Guo、Chuxuan Guo、Zijie Liu、Yiran Tong、Xuan Wu、Qining Wang 和 Caihua Xiong。用于游泳模式识别的肌肉—关节特征融合与一维卷积神经网络分类器。*IEEE*《仪器与测量汇刊》，2025年。Weizheng Chen 和 Yue Jia。利用可穿戴传感器进行网球击球动作的动力链分析。*IEEE*《仪器与测量汇刊》，2025年。Reed Ferber、Allan Brett、Reginaldo K Fukuchi、Blayne Hettinga 和 Sean T Osis。包含1,798名健康及受伤受试者在跑步机上行走与跑步期间数据的生物力学数据集。《科学数据》，11(1):1232，2024年。Sean T Osis、Blayne A Hettinga、Jessica Leitch 和 Reed Ferber。利用关节角度的主成分分析预测跑步中的足部着地时机，且不受着地方式影响。《生物力学杂志》，47 (11):2786–2789，2014年。Gregory Shakhnarovich、Trevor Darrell 和 Piotr Indyk。学习与视觉中的近邻方法。*IEEE*《神经网络汇刊》，19(2):377，2008年。Gareth James、Daniela Witten、Trevor Hastie 和 Robert Tibshirani。《统计学习导论》。Springer New York，2013年。ISBN 9781461471387。doi:10.1007/978-1-4614-7138-7。URL http://dx.doi.org/10.1007 /978-1-4614-7138-7。Bernhard Schölkopf、Alexander J Smola、Francis Bach 等人。《核方法学习：支持向量机、正则化、优化及其他》。麻省理工学院出版社，2002年。Carl Edward Rasmussen。机器学习中的高斯过程。载于机器学习暑期学校，第63–71页。Springer，2003年。Lior Rokach 和 Oded Maimon。决策树。载于《数据挖掘与知识发现手册》，第165–192页。Springer，2005年。Artur J Ferreira 和 Mário AT Figueiredo。提升算法：方法、理论与应用综述。《集成机器学习》，第35–85页，2012年。Sergio González、Salvador García、Javier Del Ser、Lior Rokach 和 Francisco Herrera。面向机器学习的装袋法与提升法集成方法实用教程：算法、软件工具、性能研究、实践观点与机遇。《信息融合》，64:205–237，2020年。Leo Breiman。随机森林。机器学习，45(1):5–32，2001年。Christopher M Bishop 等人。模式识别神经网络。牛津大学出版社，1995年。Yann LeCun、Yoshua Bengio 和 Geoffrey Hinton。深度学习。《自然》，521(7553):436–444，2015年。

Christian Szegedy、Sergey Ioffe、Vincent Vanhoucke 和 Alexander A. Alemi。Inception-v4、Inception-ResNet 以及残差连接对学习的影响。载于 第三十一届 *AAAI* 人工智能会议论文集，AAAI'17，第 4278–4284 页。AAAI Press，2017年。Jie Hu、Li Shen 和 Gang Sun。挤压激励网络。载于 *2018 IEEE/CVF* 计算机视觉与模式识别会议，第 7132–7141 页，2018年。doi:10.1109/CVPR.2018.00745。Sepp Hochreiter 和 Jürgen Schmidhuber。长短期记忆。*NeuralComputation*，9(8)：1735–1780，1997 年 11 月。ISSN 1530-888 X。doi:10.1162/neco.1997.9.8.1735。URL http://dx.doi.org/10.1162/neco.1997.9.8 .1735。Erik Štrumbelj 和 Igor Kononenko。通过特征贡献解释预测模型和个体预测。*Knowledge andInformation Systems*，41(3)：647–665，2013年8月。ISSN 0219-3116。doi:10.1007/s10115-013-0679- x。URL http://dx.doi.org/10.1007/s10115-013-0679-x。Karen Simonyan、Andrea Vedaldi 和 Andrew Zisserman。深入卷积网络内部：可视化图像分类模型和显著性图。*arXiv* 预印本 *arXiv:1312.6034*，2013年。Ramprasaath R. Selvaraju、Michael Cogswell、Abhishek Das、Ramakrishna Vedantam、Devi Parikh 和 Dhruv Batra。Grad-CAM：通过基于梯度的定位实现深度网络的视觉解释。载于 *2017 IEEE* 国际计算机视觉会议（*ICCV*），第 618–626 页。IEEE，2017年10月。doi:10.1109/iccv.2017.74。URL http://dx.doi.org/10.1109/ICCV.2017.74。Hans Van Hans Van Eetvelde、Luciana D. Mendonça、Christophe Ley、Romain Seil 和 Thomas Tischer。运动损伤预测与预防中的机器学习方法：系统综述。*Journalof Experimental Orthopaedics*，8（27），2021年。Djordje Slijepcevic、Fabian Horst、Sebastian Lapuschkin、Brian Horsak、Anna-Maria Raberger、Andreas Kranzl、Wojciech Samek、Christian Breiteneder、Wolfgang Immanuel Schöllhorn 和 Matthias Zeppelzauer。解释用于临床步态分析的机器学习模型。*ACM*医疗保健计算汇刊，3(2)，2021年。doi:10.1145/3474121。URL https://doi.org/10.1145/3474121。