**复习**

**用环境智能照亮医疗的黑暗空间**

https://doi.org/10.1038/s41586-020-2669-y✉**阿尔伯特·哈克、阿诺德·米尔施泰因和李菲菲****[[1]](" \l "_ftn1" \o ")**

# 收到日期：2020年2月16日

**接受日期：2020年7月14日在线发布日期：2020年9月9日**

机器学习和非接触式传感器的发展，使得环境智能物理空间对人类的存在非常敏感和敏感。这里我们回顾一下这项技术是如何提高我们对阴暗的，未被观察到的医疗空间。在医院空间，早期的应用可以很快实现更有效的临床工作流程，并改善重症监护室和手术室的患者安全。在常生活空间，环境智能可以延长老年人的独立性，并通过了解日常行为改善对慢性病患者的管理。与其他技术类似，大规模转化为临床应用必须克服严格的临床验证、适当的数据隐私和模型透明度等挑战。深思熟虑地使用这项技术将使我们能够了解物理环境与对健康至关重要的人类行为之间的复杂相互作用。

在数据科学和人工智能创新的推动下，决策支持系统开始帮助临床医生纠正不理想的，有时甚至是危险的诊断和治疗决策。相比之下，将更好的决定转化为临床医生、患者和家属的实际行动在很大程度上仍然是不需要帮助的。健康关键活动发生在物理空间，包括医院和私人住宅，仍然是模糊的。要获得医学进步的全部红利，部分需要不断强调负担得起的、以人为中心的方法，以帮助临床医生在这些暗喻的黑暗空间中工作。

尽管美国国家卫生研究院（NIH）、疾病控制和预防中心（CDC）、世界卫生组织（WHO）和私人组织采取了许多改进措施，例如手术安全检查表，但美国每年仍有多达40万人死于临床决策失误和缺陷和身体动作。类似的可预防的痛苦也发生在其他国家，积极性很高的临床医生也在努力应对现代医疗保健迅速增长的复杂性。为了避免压倒临床医生的认知能力，人工智能的进步有希望帮助临床医生，不仅在临床决策，而且在临床决策的物理步骤。789,106

机器学习和低成本传感器的进步可以通过提供对医疗保健的物理活动的计算机辅助理解来补充现有的临床决策支持系统。嵌入环境中的无源非接触式传感器（图1）可以形成一种环境智能，能够感知人们的运动并适应他们持续的健康需求。与现代驾驶员辅助系统类似，这种形式的环境智能可以帮助临床医生和家庭护理人员完善身体运动，这是现代医疗保健的最后一步。临床物理行动支持已经能够实现更好的制造、更安全的自主车辆和更智能的运动娱乐，能够更可靠地将生物医学发现的快速流动转化为无差错的医疗保健服务和全球人类利益。11–1415

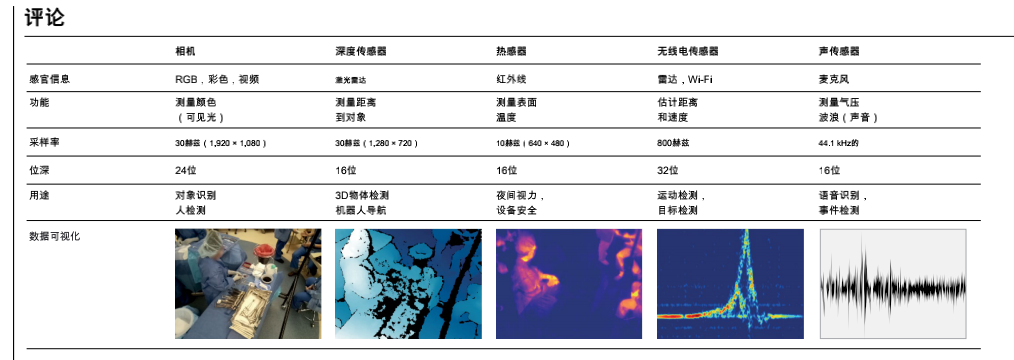
这篇综述探讨了除了基于接触的可穿戴设备外，环境非接触传感器如何照亮两个健康关键环境：医院和日常生活空间。通过几个说明性的临床用例，我们回顾了最近的算法研究和临床验证研究，列举了关键的患者结果和技术挑战。最后，我们讨论了更广泛的社会和道德考虑，包括隐私、公平、透明和道德。其他参考资料见补充说明1。

# 医院空间

2018年，约7.4%的美国人口需要住院过夜。同年，英国国家卫生局（NHS）报告了1700万次入院事件。然而，医护人员经常工作过度，医院人手不足，资源有限。我们讨论了一些医院空间，在这些空间中，环境智能可能在提高医疗服务质量、临床医生的生产力和业务运营方面发挥重要作用（图2）。这些改进在医疗危机期间可能会有很大的帮助，例如在大流行期间，医院会遇到大量的病人。161718,1920

数据可视化

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **照相机** | **深度传感器** | **热传感器** | **无线电传感器** | **声传感器** |
| 感官信息 | RGB、彩色、视频 | 激光雷达 | 红外线的 | 雷达、Wi-Fi | 话筒 |
| 功能 | 测量颜色  （可见光） | 测量到对象的距离 | 测量表面温度 | 估计距离和速度 | 测量气压波（声音） |
| 采样率 | 30赫兹（1920×1080） | 30赫兹（1280×720） | 10赫兹（640×480） | 800赫兹 | 44.1千赫 |
| 钻头深度 | 24位 | 16位 | 16位 | 32位 | 16位 |
| 使用 | 目标识别，人员检测 | 三维目标检测、机器人导航 | 夜视、设备安全 | 运动检测，目标检测 | 语音识别、事件检测 |

另一**图1 |用于环境智能的非接触式传感器。**颜色鲜艳的像素表示距离深度传感器较近的物体。黑色像素表示由反射的金属物体引起的传感器噪声。无线电传感器显示移动物体的微多普勒信号，x轴表示时间（5s），y轴表示多普勒频率。无线电传感器图像是从ref。声传感器显示讲话人的音频波形，其中x轴表示时间（5s），y轴表示信号幅度。89

## 重症监护室

重症监护室（icu）是一个专门的医院科室，在这里治疗有生命危险的疾病或严重器官衰竭的患者。在美国，icu每年花费卫生系统1080亿美元，占所有医院费用的13%。2122icu环境智能的一个很有前景的应用是计算机辅助监测病人活动。重症监护病房获得性弱点是一种常见的神经肌肉损伤的危重病人可能导致一年死亡率增加两倍，住院费用增加30%。早期患者动员可使ICU获得性弱视的相对发生率降低40%。目前，标准的流动性评估是通过直接，亲自观察，虽然它的使用受到成本不切实际，观察员偏见和人为错误的限制。正确的测量要求对患者的运动有细微的了解。例如，局部可穿戴设备可以检测步行前的动作（例如，从坐到站的过渡），但无法检测外部协助或与物理空间的交互（例如，坐在椅子上与床上）。非接触式环境传感器可以提供准确测量ICU中患者活动所需的连续和细微的理解。232425262727

在一项开创性的研究中，研究人员在一个重症监护室（图2b）安装了环境传感器（图2a），并从8名患者身上收集了362小时的数据。一个机器学习算法分类在床上，下床和步行活动的准确率为87%相比，由三名医生回顾性审查。在另一家医院进行的一项更大的研究中（图2c），另一个研究小组在八个重症监护室安装了深度传感器。他们在379个视频上训练了一个卷积神经网络，将移动活动分为四类（图2d）。在184个视频的样本外数据集上进行了验证，结果表明该算法具有87%的敏感性和89%的特异性。虽然这些初步结果是有希望的，但更深入的评估可以提供分层的结果，而不是短的、孤立的视频剪辑的总体性能。例如，一项研究使用照相机、麦克风和加速计对22名ICU患者进行了7天的监测，包括有无谵妄。研究发现，精神错乱患者的头部活动明显少于非精神错乱患者。未来的研究可以利用这项技术更快地发现谵妄，让研究人员更深入地了解患者活动如何影响死亡率、住院时间和患者康复。2829130

个早期应用是控制医院感染。在世界范围内，每年有1亿多患者受到医院获得性（即医院内）感染的影响，在重症监护室中，多达30%的患者受到医院内感染。正确遵守手卫生规程是降低医院感染频率的最有效方法之一。然而，衡量合规性仍然具有挑战性。目前，医院依靠审计员来衡量依从性，尽管费用昂贵、不连续和有偏见。可穿戴设备，特别是射频识别（RFID）徽章，是一个潜在的解决方案。不幸的是，RFID提供了粗略的位置估计（也就是说，在几十厘米之内），使得它无法对诸如WHO的手卫生五个时刻之类的细粒度运动进行分类。或者，环境传感器可以更高保真地监测洗手活动，区分酒精凝胶分配器的真实使用和临床医生在分配器附近行走。在一项开创性的研究中，研究人员在整个医院单元的壁挂式分配器上方安装了深度传感器。深度学习算法在一小时内测量351个洗手事件的依从性时达到75%的准确率。在同一时间段内，现场观察者的准确率为63%，而邻近算法（例如RFID）的准确率仅为18%。在更细致入微的研究中，环境智能检测到接触预防设备的使用以及与患者的身体接触。关键的下一步是将环境观察转化为临床行为的改变，目标是改善患者的预后。31323334353637,383940

## 手术室

在世界范围内，每年有超过2.3亿的外科手术进行，高达14%的患者经历过不良事件。这一比例可以通过更快的手术反馈来降低，例如更频繁的技术技能指导，这可以减少50%的错误。目前，外科医生的技能评估是由同行和主管，尽管是耗时，不经常和主观的。可穿戴传感器可以连接到手或仪器上，以评估外科医生的技能，但可能会抑制手的灵巧性或引入消毒的复杂性。环境摄像机是一个不引人注目的选择。一项研究训练了一个卷积神经网络来跟踪前列腺切除术视频中的针驱动器。该算法以同行评价为参考标准，将12名外科医生分为高技能组和低技能组，准确率为92%。另一项研究使用了10个胆囊切除手术的视频来重建手术过程中器械的轨迹，并将其与专家外科医生的技术评分联系起来。进一步的研究，如基于视频的手术阶段识别，可能会导致改善手术训练。然而，需要额外的临床验证，必须测试适当的反馈机制。4142434445461474849

在手术室，环境智能不仅仅限于内窥镜视频。另一个例子是手术计数-一个计数用过的物体的过程，以防止物体意外地留在病人体内。目前，需要有专门的工作人员花时间和精力对这些物体进行视觉和口头计数。由于注意力不足和团队沟通不足，人类判定的计数有可能错误地将一个物体标记为已返回505152

|  |  |  |
| --- | --- | --- |
| |  |  | | --- | --- | | 上床睡觉（两名工作人员） | 上床睡觉（一名工作人员） |   基本事实    时间09:00 09:01 09:02 |

**图2 |医院环境智能。**a、 商用环境传感器，其覆盖区域显示为绿色（即，视觉传感器的视野以及声学和无线电传感器的范围）。b、 部署在病房内的传感器可以捕捉病人、临床医生和来访者的谈话和身体运动。c、 传感器可以部署在整个医院。

**d级**深度传感器数据的活动预测和地面真相的比较。顶部，来自深度传感器的数据。中间，预测动员活动的算法、持续时间和协助病人的人数。

下面是回顾性视频回顾中人类注释的基本事实。d、 改编自参考文献。29

当它真的不见了。特别是自动计数系统，可以帮助外科团队。一项研究表明，配备条形码的剖腹手术海绵将保留物体的比率从每16天一次降低到每69天一次。RFID和Raytec海绵也有类似的结果。然而，由于其尺寸，条形码和RFID不能应用于针头和仪器，这是高达55%的计数差异的原因，每个差异平均延迟13分钟。除了海绵，环境摄像机还可以计算这些较小的物体和可能的工作人员。在一间手术室里，研究人员用天花板上安装的摄像头跟踪手术小组成员的身体部位，误差低至5厘米。在整个房间收集的环境数据可以创建术中活动的精细日志。尽管这些研究有希望作为概念的证明，但还需要进一步的研究来量化对患者结果、补偿和效率收益的影响。515354555151565758

## 其他医疗空间

临床医生将高达35%的时间花在医疗文档工作上，从而将宝贵的时间从患者身上带走。目前，医生在每次患者就诊期间或之后进行记录。一些医疗机构使用医疗记录员来减轻这一负担，导致每小时多看0.17个病人，每名病人多看0.21个相对价值单位（即保险公司报销）。然而，抄写员的培训费用很高，周转率也很高。环境麦克风可以执行与医学抄写员类似的任务。医学听写软件是另一种选择，但传统上仅限于访问后报告。在一项研究中，研究人员从90000名患者和医生之间的对话中，对14000小时的门诊音频进行深度学习模型训练。该模型显示了80%的单词级转录准确率，这表明它可能比医学抄写员的76%的准确率更好。在临床应用方面，一家医疗机构发现，戴在眼镜上的麦克风将记录时间从2小时减少到15分钟，并使与患者相处的时间增加了一倍。5960616263646562

从管理的角度来看，环境智能可以改进向作业成本法的过渡。传统上，保险公司和医院管理者通过自上而下的基于价值的会计方法估计每美元支出的健康结果。时间驱动的作业成本法是一种自下而上的替代方法，通过个人资源时间和成本（例如，使用ICU呼吸机48小时）来估计成本。这可以更好地为流程重新设计提供信息，对于一个提供者来说，流程重新设计导致患者就诊增加19%，员工减少17%，而不会降低患者结果。目前，通过亲自观察、工作人员访谈和电子健康记录，将临床活动与成本挂钩。如本综述所述，环境智能可以自动识别临床活动，统计医护人员，并估计活动的持续时间（图2d）。然而，目前缺乏环境智能的临床益处的证据，因为作业成本法对医院工作人员来说相对较新。随着技术的发展，我们希望医院管理者能够参与到环境作业成本核算系统的实施和验证中来。666768666968702929

# 日常生活空间

人类大部分时间都呆在家里。全世界的人口都在老龄化。这不仅会增加在家的时间，也会增加老年人在日常生活空间中独立生活、慢性病管理、身体康复和心理健康的重要性。71

## 老年居住空间与老龄化

到2050年，世界65岁以上人口将从7亿增加到15亿。日常生活活动，如洗澡、穿衣和进食，对这一群体的福祉和独立性至关重要。一个人的日常生活能力受损与跌倒风险增加两倍和一年死亡率增加五倍有关。早期发现损伤可提供及时临床护理的机会，有可能将日常生活能力提高两倍。目前，ADL的测量是通过自我报告的问卷调查或由护理人员手动分级，尽管这些测量是不经常的，有偏见的和主观的。另外，可穿戴设备（如加速计或心电图传感器）不仅可以跟踪日常生活能力，还可以跟踪心率、血糖水平和呼吸频率。然而，可穿戴设备无法辨别患者是否接受了ADL帮助，这是ADL评估的一个关键组成部分。非接触式环境传感器（图3a）可以潜在地识别这些临床细微差别，同时检测更大范围的活动。717273117475767778

在第一个此类研究中，研究人员在一个老年人的卧室内安装了一个深度和温度传感器（图3b），并在一个月内观察了1690项活动，包括231例护理人员协助（图3c）。卷积神经网络检测辅助的准确率为86%。在另一项研究中，研究人员从一个老人院的6个人那里收集了10天的视频，并取得了类似的结果。尽管视觉传感器很有前途，但在某些环境中，它们会引起隐私问题，例如浴室，那里是梳洗、洗澡和如厕活动发生的地方，所有这些都强烈地表明了认知功能。这使得研究人员探索声学和雷达传感器。一项研究使用麦克风检测淋浴和如厕活动，准确率分别为93%和91%。然而，这些研究的一个局限性是它们在少数环境中的评价。日常生活空间是高度可变的，因此带来了普遍化的挑战。此外，隐私是最重要的。如果这项技术要照亮日常生活空间，开发和验证安全、隐私安全的系统是必不可少的。7918081828382

老年人独立生活的另一个应用是跌倒检测。大约29%的社区成年人每年至少跌倒一次。跌倒后躺在地上超过一小时与12个月死亡率增加5倍相关。此外，由于跌倒检测系统的安全性，与抑郁和低生活质量相关的跌倒恐惧可以减少。几十年来，研究人员开发了带有可穿戴设备和非接触式环境传感器的坠落探测系统。一项系统研究发现，可穿戴设备检测跌倒的准确率为96%，而环境传感器的准确率为97%。在另一项研究中，研究人员在271个家庭中安装了蓝牙（即无线电）信标。利用每个信标的信号强度，机器学习算法对老年人的弱点进行分类，准确率为98%。在另一项研究中，研究人员在16个老年公寓的天花板上安装了深度和雷达传感器，持续了2年。通过小波分解变换后的雷达信号，可以检测到100%的坠落，每天不到两次假警报。深度传感器每月产生一次假警报，坠落探测率为98%。环境传感器的速度足够快（即低延迟），可以向13个辅助生活社区的护理人员提供实时电子邮件警报。与一年以上85名老年人组成的对照组相比，实时干预显著减缓了86名老年人的功能衰退。当与可穿戴设备结合使用时，一项研究发现，深度传感器的坠落检测准确率从90%提高到98%，表明非接触式和可穿戴式传感器之间可能存在协同效应。随着环境智能开始弥合观察和干预之间的鸿沟，需要进一步的研究来探索监管批准流程、法律含义和道德考虑。84858687888990919293949596

## 慢性病管理

步态分析应用于身体康复和慢性疾病，是诊断测试和测量治疗效果的重要工具。例如，频繁而准确的步态分析可以改善脑瘫儿童术后的健康状况，或使帕金森氏病的早期发现可以长达4年。传统上，步态分析仅限于具有力板和运动捕捉系统的研究实验室，越来越多地使用可穿戴设备进行步态分析。一项研究使用加速计来估计30名慢性肺病患者6分钟的临床标准步行距离。研究发现平均绝对误差率为6%。一个限制是可穿戴设备必须与身体相连，给患者带来不便。另外，非接触式传感器可以持续测量步态，提高保真度，并创建交互式的，基于家庭的康复计划。有几项研究用照相机、深度传感器、雷达和麦克风在自然环境下测量步态。一项研究使用深度传感器测量了9名帕金森病患者的步态模式。利用高端运动捕捉系统作为地面真实情况，研究发现，深度传感器可以跟踪膝盖垂直运动到4厘米以内。另一项研究利用深度传感器为脑瘫患者制作了一个运动游戏。根据Tinetti测试，在24周的时间里，使用游戏的患者平衡和步态提高了18%。虽然有希望，这些研究评估了单一传感器模式。在实验室实验中，当麦克风和可穿戴传感器相结合时，步态检测提高了3%到7%。在可行的情况下，研究可以调查多种传感方式（如被动红外运动传感器、接触式传感器和可穿戴相机）的潜在协同效应。979899100101102103104105104106107108109110111

## 心理健康

抑郁症、焦虑症和双相情感障碍等精神疾病影响着美国4300万成年人和欧盟1.65亿人。据估计，56%的患有精神疾病的成年人由于经济成本和医疗机构的可用性等障碍而不寻求治疗。目前，自我报告的问卷调查和临床评估（例如，精神障碍诊断和统计手册（DSM-5））是识别精神疾病症状的标准工具，尽管很少出现和有偏见。或者，环境传感器可以提供连续和成本效益的症状筛查。在一项研究中，研究人员在30分钟的半结构化临床访谈中收集了69个人的音频、视频和深度数据。利用患者的言语线索和上身运动，机器学习算法检测出46例精神分裂症患者，阳性预测值为95%，敏感性为84%。同样，在急诊科，通过对61名青少年（其中31名自杀者）的临床访谈进行自然语言分析，得出了一个模型，该模型能够对自杀患者进行90%的分类112113112114115116

|  |
| --- |
|  |
|  |  |

**图3 |日常生活空间的环境智能。**a、 老人之家配有一个环境传感器。绿色平截头体表示传感器的覆盖区域（即，视觉传感器的视野以及声学和无线电传感器的范围）。b、 来自传感器的热量和深度数据由环境智能算法处理，用于活动分类。c、 患者一天活动的总结。深蓝色部分表示活动更频繁。c、 改编自参考文献。79

准确度。虽然令人印象深刻，但还需要进一步的试验来验证对患者结局的影响。117

然而，即使在被发现后，治疗精神健康疾病仍然很复杂。特殊的治疗师效应会导致高达17%的结果差异，这使得进行心理治疗研究变得困难。转录本是识别良好治疗特征的标准方法，但收集起来很昂贵。人工编码一个20分钟的疗程可以从85分钟到120分钟不等。环境传感器可以为心理治疗研究提供更便宜、更高质量的成绩单。其中一项研究使用了一种递归神经网络，利用治疗过程中的短信，从14899名患者身上检测出24种治疗师技术的实例。该研究确定了几种与改善患者健康问卷（PHQ-9）和一般性焦虑症（GAD-7）评分相关的技术。另一项研究使用麦克风和语音识别算法来转录和评估200次20分钟动机式访谈中治疗师的移情。以人类评估委员会为金标准，该算法的准确率为82%。尽管这比单个人类评估员90%的准确率要低，但是环境智能可以更容易地应用于更多的患者。利用环境智力，研究人员现在可以进行大规模的研究，以重申他们对心理治疗框架的理解。然而，还需要进一步的研究来验证这些系统对不同人群的治疗师和患者的普遍性。1181191201121120120

# 技术挑战和机遇

通过观察与康复相关的行为、减少临床医生的意外失误、协助老龄化人口和监测慢性病患者，环境智能有可能照亮医疗保健提供过程。在表1中，我们强调了与在复杂场景中识别人类行为以及从临床环境中的大数据和罕见事件中学习有关的七项技术挑战和机遇。

## 复杂场景下的行为识别

了解医疗领域中复杂的人类行为需要跨机器智能多个领域的研究，如视觉跟踪、人体姿势估计和人-物交互模型。考虑一下在医院晨诊。多达十几个临床医生系统地审查和访问每个病人在医院单位。在此期间，临床医生可能会遮挡传感器对患者的观察，从而可能导致对健康至关重要的活动无法被发现。如果一个物体在遮挡前移动，跟踪算法（表1）可以在遮挡时估计物体的位置。对于较长的遮挡，矩阵完成方法（如图像修复）可以“填充”遮挡对象背后的内容。类似的技术可用于频谱图形式的音频去噪。如果没有阻塞，下一步就是找到人。在晨诊期间，临床医生可能会互相递上物品或指点房间，从传感器的角度引入身体的多层部分。人体姿态估计算法（表1）试图通过精确定位身体部位并将其分配给正确的个体来解决这种模糊性。建立高度精确的人类行为模型是环境智能在复杂的临床环境中取得成功的必要条件。122123124125

环境智能需要了解人类如何与物体和其他人互动。一类方法试图识别图像中的视觉基础关系，通常以126

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **表1 |算法挑战**   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | 挑战子挑战 | | 技术方法 | 重症监护病房 | | 手术室 | | | 其他 | | 老年人护理 | | 慢性 | 心理健康 | | | 患者移动性 | 手卫生 | 技能 | 手术计数 | 笔记 | | 成本计算 | 日常生活用品 | 瀑布 | 步态分析 | 症状治疗筛选研究 | | | 复杂环境中的行为复杂识别环境 | | 视觉跟踪，矩阵完成 | 十 | 十 | 十 | 十 | − | | 十 | 十 | 十 | 十 | 十 | − | | 场景 | 定位多人 | 行人检测、人体姿态估计 | 十 | − | − | − | − | | 十 | 十 | 十 | 十 | − | − | | 认识人类行为 | 场景图，活动识别 | 十 | 十 | − | 十 | − | | 十 | 十 | 十 | − | − | 十 | | 利用大数据和罕见事件进行学习 | 大数据 | 分布式学习，优化器 | 十 | 十 | 十 | 十 | 十 | | 十 | 十 | 十 | 十 | 十 | 十 | | 实时检测 | 两阶段模型，模型压缩 | − | 十 | 十 | 十 | − | | − | − | 十 | − | − | − | | 罕见事件 | 校准，损失加权 | 十 | − | − | 十 | − | | − | − | 十 | − | 十 | − | | 泛化迁移  对于新的学习，很少有机会  学习 | | 十 | 十 | 十 | 十 | 十 | | 十 | 十 | 十 | 十 | 十 | 十 | |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |   行表示算法挑战。列表示临床用例。适用于特定临床用例的挑战以“x”标记“技能”是指对外科手术的评价 |

|  |
| --- |
|  |
|  |  |

技能；“注释”指的是医疗文件。

场景图（表1）。场景图是由相互连接的节点组成的网络，其中每个节点表示图像中的一个对象，每个连接表示它们之间的关系。场景图不仅可以帮助识别人类行为，还可以使环境智能更加透明。127128

## 从大数据和罕见事件中学习

环境传感器将从医院和家庭产生数PB的数据。这就需要新的机器学习方法，这些方法能够对罕见事件进行建模，并处理待开发的大数据（表1）。大规模的活动理解模型可能需要数天的训练，除非使用大型集群的专用硬件。云服务器是一个潜在的解决方案，但可能会很昂贵，因为环境智能可能需要大量的存储、计算和网络带宽。改进的基于梯度的优化器和神经网络结构可以潜在地减少训练时间。然而，快速训练一个模型并不能保证它在推理过程中是快速的（即实时检测）（表1）。例如，基于视频的活动识别模型速度较慢，通常为每秒1到10帧。即使是每秒100帧的优化模型，每天处理数兆字节的数据也可能有困难。模型压缩和量化等技术可以减少存储和计算需求。有些方法不是以完全的空间或时间分辨率处理音频或视频，而是快速识别感兴趣的片段，称为建议。这些建议，然后提供给重型模块高度准确，但计算密集的活动识别。129130131132133134135136137

尽管环境传感器产生的数据量很大，但一些临床事件很少发生（表1）。这些长尾事件的检测对于了解健康关键行为是必要的。以坠落检测为例。大多数环境数据包含正常活动，由于标签不平衡，使算法产生偏差。更广泛地说，统计偏差可以应用于任何类别的数据，例如受保护的类属性。一种解决方案是对算法进行统计校准，从而使指定属性的错误率保持一致。然而，一些医疗环境中跌倒的发生率可能比原来的训练环境中要高。这需要泛化（表1）：一个算法对看不见的分布进行操作的能力。代替训练一个为所有分布设计的模型，另一种方法是采用一个现有的模型，并在新的分布（也称为转移学习）上对其进行微调。另一个解决方案，域自适应，试图通过更好的特征表示来最小化训练和测试分布之间的差距。对于低资源的医疗保健提供者，很少有能够从一两个例子中学习的镜头学习算法可以使用。138139140141142143144

# 社会和道德考虑

环境智能系统的可靠性是实现这项技术潜力的关键。尽管关于可信人工智能的文献越来越多，但我们考虑了可信的四个独立维度：隐私、公平、透明和研究伦理。在解决这四个因素的同时发展这项技术需要来自医学、计算机科学、法律、伦理和公共政策的专家之间的密切合作。145

## 隐私

通过设计，环境传感器可以持续观察环境，并可以发现有关人类身体行为如何影响医疗服务提供的新信息。例如，传感器可以远距离测量生命体征。虽然方便，这些知识可能被用来推断私人医疗条件。随着全球公民对海量数据收集越来越敏感，人们对这些信息的保密性、共享和保留越来越关注。因此，必须在考虑隐私和安全的情况下共同开发这项技术，不仅要考虑到技术本身，而且要考虑到所有利益相关者在开发过程中的持续参与。146147148

图4中展示了许多现有的和新兴的隐私保护技术。一种方法是通过删除个人的身份来取消对数据的识别。另一种方法是数据最小化，

**计算**

### 方法描述硬件转换结果

|  |  |  |
| --- | --- | --- |
| 差别隐私 | 为数据添加噪波；影响最小  人口水平分析 | 边缘计算机 |
| 面部模糊 | 检测并模糊人脸 | 传感器、边缘计算机 |
| 降维 | 通过减少功能的数量来减少输入大小 | 传感器、边缘计算机 |
| 身体掩蔽 | 用不要脸的化身代替人 | 边缘计算机 |
| 联合学习 | 边缘设备在本地学习，然后向中心服务器发送梯度更新 | 边缘  计算机、集中式服务器 |
| 同态加密 | 允许从加密数据进行预测 | 边缘  计算机、集中式服务器 |

**图4 |保护隐私的计算方法。**在每个方法提供的隐私保护级别和所需的计算资源之间存在权衡。用于生成变换图像的方法在别处详细描述：差分隐私，参考文献；降维，参考文献；身体掩蔽，参考文献；联合学习，参考文献；同态加密，参考文献。原图由S.McCoy制作，此前已经出版。美国国防部可视信息的出现并不意味着或构成美国国防部的认可。166167168169170171

最大限度地减少了数据采集、传输和人为副渔获物。当病人没有占用病房时，环境系统可能会暂停。然而，即使数据被取消识别，也有可能重新识别个人。超分辨率技术可以部分逆转人脸模糊和降维技术的影响，有可能实现重新识别。这表明数据应该留在设备上，以减少未经授权的访问和重新识别的风险。149150

法律和社会的复杂性将不可避免地出现。有记录在案的例子表明，要求公司向执法部门提供来自环境扬声器和摄像头的数据。尽管这些装置位于潜在的犯罪现场内，但这就提出了一个问题，即在什么时候应该披露犯罪现场外的偶然发现，例如无意中的供词。与数据共享相关的是，一些医疗保健组织与第三方（如数据代理）共享患者信息。为了缓解这种情况，患者应主动要求医疗保健提供者使用隐私保护措施（图4）。此外，临床医生和技术人员必须与关键利益相关者（例如，患者、家属或护理人员）、法律专家和决策者合作，为环境系统开发治理框架。151152

## 公平

环境智能将与大量患者群体相互作用，可能比当前临床医生的影响范围大几个数量级。这迫使我们审视环境制度的公平性。公平是一个复杂而多方面的话题，多个研究团体都在讨论。我们在这里强调算法公平性的两个方面作为例子：数据集偏差和模型性能。138

标记数据集是大多数机器学习系统的基础。然而，医学数据集甚至在深入学习之前就存在偏见。这些偏见会对某些人群的临床结果产生不利影响。如果一个人缺少特定的属性，无论是由于数据收集的限制还是社会因素，算法可能会误解他们的整个记录，导致更高水平的预测错误。识别偏差的一种方法是分析不同组的模型性能。在一项研究中，当预测30天的精神病再入院率时，不同种族的错误率不同。更严格的方法可以检验等灵敏度和等阳性预测值。然而，相同的模型性能可能不会产生相同的临床结果，因为某些人群可能有内在的生理差异。尽管如此，在减少偏见方面正在取得进展，例如先证者工具。1153154155156157158

## 透明度

环境智能可以揭示人类行为对医疗服务提供的影响。这些发现可能会让一些研究人员感到惊讶，在这种情况下，临床医生和患者在使用前需要相信这些发现。环境智能系统应提供可解释的预测性、描述性和相关性结果，而不是不透明的黑盒模型。这有助于获得利益相关者的认可，因为技术文盲和模型不透明会阻碍在医疗保健中使用环境智能的努力。透明度并不局限于算法。数据集透明度——详细跟踪数据集的设计、收集和注释方式，以便为将来的应用采取具体的预防措施，例如培训人类注释者或修订研究的纳入和排除标准。关于透明度的正式准则，如三脚架声明，正在积极制定之中。另一个工具是使用模型卡，这是一种简短的分析，在不同人群中对算法进行基准测试，并概述评估程序。159160161162

## 研究伦理学

伦理研究包括保护人类参与者、独立审查和公共福利等主题。贝尔蒙特的报告促使对涉及人类参与者的研究进行监管，其中将“尊重人”作为一项基本原则。在研究中，这表现为研究参与者的知情同意。然而，如果研究对参与者造成的风险最小或无法获得同意，一些法规允许在未经同意的情况下进行研究。对于大规模的环境智力研究，获得知情同意可能很困难，而且在某些情况下，由于自动去识别技术，这可能是不可能的（图4）。在这些情况下，公众参与或协商民主可以是替代的解决办法。163

仅仅依靠主要调查人员的诚信来进行伦理研究可能会带来潜在的利益冲突。为了降低这种风险，涉及人类参与者的学术研究需要获得机构审查委员会的批准。公共卫生监督旨在防止疾病蔓延和改善健康，不需要独立审查。根据应用的不同，环境智能可分为两类。研究人员被敦促咨询法律和伦理专家，以确定适当的步骤来保护所有人类参与者，同时最大限度地提高公共福利。164165

# 摘要

几个世纪的医学实践导致了知识爆炸，推动了人类健康的空前进步。人工智能和低成本非接触式传感器的突破性进展催生了一种环境智能，它有可能改善医疗服务的物理执行。来自医院和日常生活空间的初步结果证实了通过环境感应获得的丰富信息。这一照亮医疗黑暗空间的非凡机会要求计算机科学家、临床医生和医学研究人员与法律、伦理和公共政策方面的专家密切合作，为医疗保健创造可信赖的环境智能系统。

1.       LeCun，Y.，Bengio，Y.&Hinton，G.《深度学习》。《自然》521436–444（2015）。本文回顾了深度学习的发展，并解释了常用的神经网络结构，如卷积神经网络和递归神经网络在视觉和自然语言处理任务中的应用。

2.       Jordan，M.I.&Mitchell，T.M.《机器学习：趋势、前景与展望》。《科学》349255–260（2015）。

3.       Esteva，A.等人，《医疗深度学习指南》。纳特。地中海。25, 24–29 (2019). 这一观点描述了计算机视觉、自然语言处理、语音识别和强化学习在医学成像任务、电子健康记录分析、机器人辅助手术和基因组研究中的应用。

4.       高性能医学：人类与人工智能的融合。纳特。地中海。25, 44–56 (2019).

**本文概述了临床医生、患者和卫生系统如何利用人工智能来解释医学图像、发现工作流程效率和促进患者自我护理**.

5.       临床决策支持系统概述：益处、风险和成功策略。NPJ数字。地中海。3, 17 (2020).

6.       Yeung，S.，Downing，N.L.，Fei Fei，L.&Milstein，A.。床边计算机视觉-将人工智能从驾驶员辅助转移到患者安全。N、 英语。J、 地中海。378, 1271–1273 (2018).

7.       Haynes，A.B.等人，《全球人口中降低发病率和死亡率的手术安全检查表》。N、 英语。J、 地中海。360, 491–499 (2009).

8.       Makary，M.A.&Daniel，M.。医疗失误是美国第三大死因。比尔。地中海。J、 353，i2139（2016年）。

9.       Tallerther，V.R.，Smith，S.E.，Skinner，J.和Cameron，H.S.《探索基于团队的急性护理方案中的错误：来自英国的一项观察性研究》。阿卡德。地中海。87, 792–798 (2012).

10.     杨树泰等。2002-2011年中国医疗事故诉讼评估。J、 法医学。地中海。2, 185–189 (2016).

11.     Pol，M.C.，ter Riet，G.，van Hartingsveldt，M.，Kröse，B.&Buurman，B.M.。髋部骨折后老年患者康复计划中传感器监测的有效性：一项三臂阶梯楔形随机试验。年龄在48650-657之间（2019年）。

12.     Fritz，R.L.和Dermody，G.一种护士驱动的方法，用于在“智能”养老院中开发人工智能。努尔斯。展望67、140–153（2019年）。

13.     Kaye，J.A.等人，F5-05-04：生命活动的生态有效性评估：用传感器进行不引人注目的连续监测。阿尔茨海默痴呆症。12，第374页（2016年）。

14.     Acampora，G.，Cook，D.J.，Rashidi，P.&Vasilakos，A.V.《医疗环境智能调查》。电气工程师学会。电子。工程101、2470–2494（2013年）。

15.     库克，D.J.，邓肯，G.，斯普林特，G.&弗里茨，R.。使用智能城市技术使医疗保健更加智能化。电气工程师学会。电子。英语106、708–722（2018年）。

16.     疾病控制和预防中心。全国健康访谈调查：健康统计摘要https://www.cdc.gov/nchs/nhis/shs.htm(2018).

17.     NHS数字。2018-19年住院患者护理和成人重症护理活动https://digital.nhs.uk/data-and-information/publications/statistical/住院患者护理活动/2018-19（NHS，2019）。

18.     Patel，R.S.，Bachu，R.，Adikey，A.，Malik，M.&Shah，M.。与医生职业倦怠及其后果相关的因素：综述。比哈夫。科学院。（巴塞尔）8，98（2018）。

19.     Lyon，M.等人，《因缺乏放射服务而导致的农村ED转移》。是。J、 紧急情况。地中海。33, 1630–1634 (2015).

20.     Adams，J.G.和Walls，R.M.在COVID-19全球流行期间支持卫生保健工作人员。J、 是的。地中海。协会第323、1439–1440（2020年）。

21.     Halpern，N.A.，Goldman，D.A.，Tan，K.S.和Pastores，S.M.《美国人口群体、医疗保险和医疗补助受益人的危重病护理床位和使用趋势：2000-2010》。暴击。医疗护理。44, 1490–1499 (2016).

22.     Halpern，N.A.&Pastores，S.M.《2000-2005年美国危重病护理医学：床位数量、入住率、付款人组合和成本分析》。暴击。医疗护理。38, 65–71 (2010).

23.     Hermans，G.等人，《重症监护室获得性虚弱的急性转归和1年死亡率》。队列研究和倾向匹配分析。是。J、 回应。暴击。医疗护理。190, 410–420 (2014).

24.     张，L。等。重症监护室危重病人的早期动员：系统回顾和荟萃分析。《公共科学图书馆·综合》第14期，第223185页（2019年）。

25.     Donchin，Y.等人，《重症监护病房人为失误的性质和原因研究》。暴击。医疗护理。23, 294–300 (1995).

26.     Hodgson，C.L.，Berney，S.，Harrold，M.，Saxena，M.&Bellomo，R.。临床回顾：ICU中的早期患者动员。暴击。Care 17，207（2013年）。

27.     Verceles，A.C.&Hager，E.R.利用加速度计监测危重病人的体力活动：一项系统综述。回应。Care 60，1330–1336（2015年）。

28.     Ma，A.J.等人，使用一种新型无创传感器测量ICU患者的活动度。暴击。医疗护理。45, 630–636 (2017).

29.     Yeung，S.等。ICU中基于深度学习的患者活动检测计算机视觉系统。NPJ数字。地中海。2, 11 (2019).

**这项研究使用计算机视觉同时对重症监护病房的患者动员活动进行分类，并统计参与每项活动的医护人员数量**.

30.     Davoudi，A.等。智能ICU，使用普及感知和深度学习进行患者自主监测。科学院。第98020页（2019年）。

**这项研究使用相机和可穿戴传感器来跟踪重症监护室中精神错乱和非精神错乱患者的身体运动。**

31.     谁。全球地方卫生保健相关感染负担报告https://apps.who.int/iris/handle/10665/80135(2011).

32.     文森特，J.-L。成人重症监护病房的医院感染。《柳叶刀》3612068–2077（2003）。

33.     Gould，D.J.，Moralejo，D.，Drey，N.，Chudleigh，J.H.&Taljaard，M.《患者护理中改善手卫生依从性的干预措施》。科克伦数据库系统。版次。9，CD005186（2017年）。

34.     Srigley，J.A.，Furness，C.D.，Baker，G.R.和Gardam，M.使用电子监测系统对手卫生依从性监测中的霍桑效应进行量化：一项回顾性队列研究。BMJ质量。苏丹武装部队。23, 974–980 (2014).

35.     Shirehjini，A.A.N.，Yassine，A.和Shirmohammadi，S.使用基于RFID定位系统的医院设备定位。IEEE传输。信息技术。生物群落。16, 1058–1069 (2012).

36.     Sax，H.等人，“手卫生的五个时刻”：一种以用户为中心的设计方法，用于理解、培训、监控和报告手卫生。J、 医院。感染。67, 9–21 (2007).

37.     Haque，A.等人，《走向基于视觉的智能医院：跟踪和监测手卫生合规性的系统》。进行中。第二届医疗保健机器学习会议75–87（PMLR，2017）。

**本研究评估了深度传感器和隐蔽审核员在衡量医院单位手卫生依从性方面的表现**.

38.     Singh，A.等人。利用计算机视觉技术自动检测手部卫生状况。J、 是的。地中海。通知。协会。https://doi.org/10.1093/jamia/ocaa115(2020).

39.     Chen，J.，Cremer，J.F.，Zarei，K.，Segre，A.M.和Polgreen，P.M.使用计算机视觉和深度传感测量医务人员与患者的接触情况和病房内个人防护设备的依从性。打开论坛感染。数字化信息系统。3，ofv200（2016）。

40.     Awwad，S.，Tarvade，S.，Piccardi，M.和Gattas，D.J.使用隐私保护的计算机视觉来测量医护人员手部卫生的质量。内景J.夸尔。医疗31，36–42（2019年）。

41.     Weiser，T.G.等人，《全球外科手术量的估计：基于可用数据的建模策略》。《柳叶刀》372139–144（2008）。

42.     Anderson，O.，Davis，R.，Hanna，G.B.&Vincent，C.A.《外科不良事件：系统性回顾》。是。J、 外科学。206, 253–262 (2013).

43.     Bonrath，E.M.，Dedy，N.J.，Gordon，L.E.&Grantcharov，T.P.。综合外科指导提高手术室的外科技能：一项随机对照试验。安。外科学。262, 205–212 (2015).

44.     Vaidya，A.等人，《外科技术技能评估工具的现状：系统综述》。J、 外科学。第246342–378号决议（2020年）。

45.     Ghasemloonia，A.等人，《使用运动质量和平滑度进行外科技能评估》。J、 外科学。教育。74, 295–305 (2017).

46.     Khalid，S.，Goldenberg，M.，Grantcharov，T.，Taati，B.&Rudzicz，F.《识别手术行为和测量绩效的深度学习模型评估》。JAMA网络。打开3，e201664（2020）。

47.     Law，H.，Ghani，K.&Deng，J.。基于计算机视觉分析的外科医生技术技能评估。进行中。第二届医疗保健机器学习会议88–99（PMLR，2017）。

48.     Jin，A.等。基于区域的卷积神经网络在手术视频中的工具检测和操作技能评估。进行中。计算机视觉应用冬季会议691–699（IEEE，2018）。

49.     Twinanda，A.P.等人，《内联网：腹腔镜视频识别任务的深层结构》。IEEE传输。地中海。影像36，86–97（2017）。

50.     Hashimoto，D.A.，Rosman，G.，Rus，D.&Meireles，O.R.《手术中的人工智能：承诺与危险》。安。外科学。268, 70–76 (2018).

51.     Greenberg，C.C.，Regenbogen，S.E.，Lipsitz，S.R.，Diaz Flores，R.&Gawande，A.A.。手术计数差异的频率和意义。安。外科学。248, 337–341 (2008).

52.     Agrawal，A.《计数问题：保留手术项目根本原因分析的经验教训》。日本烟草公司。通信质量。患者安全气囊。38, 566–574 (2012).

53.     Hempel，S.等人，《错误的手术部位、保留的手术物品和手术火灾：

手术事件的系统回顾。贾马苏格。150, 796–805 (2015).

54.     Cima，R.R.等人，使用数据矩阵编码海绵计数系统进行外科实践：18个月后的影响。日本烟草公司。通信质量。患者安全气囊。37, 51–58 (2011).

55.     Rupp，C.C.等人。射频检测系统作为追踪外科海绵的人工计数方案的辅助手段的有效性：2285名患者的前瞻性试验。J、 是的。科尔。外科学。215, 524–533 (2012).

56.     Kassahun，Y.等人，《增强灵巧仪器之外的手术机器人：机器学习技术及其在智能和自主手术操作中的作用的调查》。国际计算机杂志。协助。放射科医生。外科学。11, 553–568 (2016).

57.     Kadkhodammohammadi，A.，Gangi，A.，de Mathelin，M.&Padoy，N.。手术室人体姿势估计的多视角RGB-D方法。进行中。计算机视觉应用冬季会议363–372（IEEE，2017）。

58.     Jung，J.J.，Jüni，P.，Lebovic，G.&Grantcharov，T.《手术室黑匣子研究的第一年分析》。安。外科学。271, 122–127 (2020).

59.     Joukes，E.，Abu Hanna，A.，Cornet，R.&de Keizer，N.F.。在引入结构化和标准化的电子健康记录之前和之后，花费在专门的患者护理和文档工作上的时间。申请。临床。通知。9, 46–53 (2018).

60.     Heaton，H.A.，Castaneda Guarderas，A.，Trotter，E.R.，Erwin，P.J.&Bellolio，M.F.《抄写员对患者吞吐量、收入以及患者和提供者满意度的影响：系统回顾和荟萃分析》。是。J、 紧急情况。地中海。34, 2018–2028 (2016).

61.     Rich，N.医学抄写员工作的影响。是。J、 紧急情况。地中海。35, 513 (2017).

62.     Boulton，C.《谷歌眼镜如何为尊严健康自动化患者文档》。

*《华尔街日报》*（2014年6月16日）。

63.     Blackley，S.V.，Huynh，J.，Wang，L.，Korach，Z.&Zhou，L.《1990-2018年临床文档语音识别：系统性综述》。J、 是的。地中海。通知。协会第26、324–338号（2019年）。

64.     Chiu，C.-C.等人。医学对话的语音识别。进行中。国际语音通信协会第18届年会2972–2976（ISCA，2018）。

**本文提出了一种语音识别算法来转录患者和临床医生之间的匿名对话**.

65.     Pranaat，R.等人。使用基于电子健康记录环境的模拟来评估医学抄写员生成的笔记的结构和准确性：

概念验证研究。杰米尔医学院。通知。5，e30（2017）。

66.     Kaplan，R.S.等人，利用时间驱动的作业成本法来确定医疗保健领域的价值提升机会。J、 健康C。管理。59, 399–412 (2014).

67.     波特，M.E.基于价值的医疗服务提供。安。外科学。248, 503–509 (2008).

68.     Keel，G.，Savage，C.，Rafiq，M.&Mazzocato，P.《医疗保健中时间驱动的作业成本法：文献的系统回顾》。卫生政策121，755–763（2017）。

69.     French，K.E.等人，使用时间驱动的作业成本法在术前评估中心衡量过程改进计划的价值。医疗保健1，136–142（2013年）。

70.     Sánchez，D.，Tentori，M.&Favela，J.《智能医院的活动识别》。IEEE智能系统23，50–57（2008）。

71.     联合国。2019年世界人口老龄化https://www.un.org/development/desa/pd/站点/www.un.org.development.desa.pd/files/files/documents/2020/Jan/un\u 2019\世界人口老龄化\_报告.pdf(2020).

72.     Mamikonian Zarpas，A.&Laganá，L.。老年人未来跌倒风险与日常生活活动困难之间的关系。J、 衰老的老年病学。3, 8–16 (2015).

73.     Stineman，M.G.等人根据日常生活阶段的活动，都会导致老年人1年、5年和10年的死亡率。J、 是的。老年人。Soc公司。60, 485–492 (2012).

74.     Phelan，E.A.，Williams，B.，Penninx，B.W.J.H.，LoGerfo，J.P.&Leveille，S.G.健康促进计划随机试验中老年人日常生活功能和残疾的活动。J、 老年病学。A 59，M838–M843（2004年）。

75.     Carlsson，G.，Haak，M.，Nygren，C.&Iwarsson，S.。自我报告与专业评估的社区老年人功能限制。国际康复杂志。第35299-304号决议（2012年）。

76.     Wang，Z.，Yang，Z.&Dong，T.。可穿戴式老年护理技术综述，该技术能够准确跟踪室内位置，识别身体活动并实时监测生命体征。传感器17，341（2017）。

77.     评估自我维持：日常生活活动、移动性和工具性日常生活活动。J、 是的。老年人。Soc公司。31, 721–727 (1983).

78.     Uddin，M.Z.，Khaksar，W.和Torresen，J.《老年护理和独立生活环境传感器：一项调查》。传感器18，2027（2018）。

79.     Luo，Z.等。基于计算机视觉的老年人日常活动描述性分析，用于长期健康监测。进行中。第三届医疗保健机器学习会议1–18（PMLR，2018）。

**这项研究创造了空间和时间的活动总结日常生活中使用的深度和热传感器内的一个老年居民的卧室**.

80.     Cheng，H.，Liu，Z.，Zhao，Y.，Ye，G.&Sun，X.《老年家庭监护的真实世界活动总结》。多媒体工具应用。70, 177–197 (2014).

81.     Lee，M.-T.，Jang，Y.&Chang，W.-Y.认知功能受损如何影响老年人的日常生活功能活动？《公共科学图书馆·综合》第14期，第e0218112页（2019年）。

82.     Chen，J.，Zhang，J.，Kam，A.H.&Shue，L.。浴室自动声学监控系统。进行中。1750-1753年电路与系统国际研讨会（IEEE，2005）。

83.     Shrestha，A.等人，《老年护理：S波段雷达日常生活活动分类》。J、 工程2019，7601–7606（2019）。

84.     Ganz，D.A.&Latham，N.K.。社区老年人跌倒预防。N、 英语。J、 地中海。382, 734–743 (2020).

85.     Bergen，G.，Stevens，M.R.&Burns，E.R.《65岁以上成年人的跌倒和跌伤》——美国，2014年。莫布。凡人。好极了。报告65993–998（2016年）。

86.     Wild，D.，Nayak，U.S.&Isaacs，B.家中老人跌倒有多危险？比尔。地中海。J、 （临床。第28226–268页（1981年）。

87.     Scheffer，A.C.，Schuurmans，M.J.，van Dijk，N.，van der Hooft，T.&de Rooij，S.E.《老年人跌倒恐惧：测量策略、患病率、风险因素和后果》。37岁，19-24岁（2008年）。

88.     Pol，M.等人，《老年人对在家中使用传感器监测的看法》。老年学家56，485–493（2016）。

89.     Erol，B.，Amin，M.G.&Boashash，B.。用于坠落探测的距离多普勒雷达传感器融合。进行中。IEEE雷达会议819–824（IEEE，2017）。

90.     Chaudhuri，S.，Thompson，H.&Demiris，G.《跌倒检测装置及其在老年人中的应用：系统综述》。J、 老年人。物理。疗法。37, 178–196 (2014).

91.     Tegou，T.等人。一种用于检测老年人虚弱的低成本室内活动监测系统。传感器19，452（2019）。

92.     Rantz，M.等人。老年人家庭跌倒风险自动评估和检测传感器系统。老年学家55，S78–S87（2015）。

93.     Su，B.Y.，Ho，K.C.，Rantz，M.J.&Skubic，M.。使用小波变换的多普勒雷达坠落活动检测。IEEE传输。生物群落。Eng.62865–875（2015年）。

94.     Stone，E.E.&Skubic，M.使用微软Kinect在老年人家中进行跌倒检测。IEEE J.生物医学。健康信息。19, 290–301 (2015).

95.     Rantz，M.等人。智能传感器系统用于老年住宅早期疾病警报的随机试验。J、 是的。地中海。目录。协会第18，860–870号（2017年）。

**这项随机试验研究了由异常步态触发的实时干预系统的临床疗效，该系统由环境传感器对家中老年人步行能力的检测**.

96.     Kwolek，B.&Kepski，M.。使用深度图和无线加速度计在嵌入式平台上进行人体坠落检测。计算机。方法：程序生物医学。117, 489–501 (2014).

97.     Wren，T.A.L.，Gorton，G.E.III，Ounpuu，S.&Tucker，C.A.临床步态分析的疗效：系统评价。步态姿态34149-153（2011）。

98.     Wren，T.A.等。有步态分析和无步态分析的脑瘫患儿下肢矫形手术的结果：一项随机对照试验的结果。步态姿势38236–241（2013）。

99.     Del Din，S.等人。可穿戴设备步态分析预测帕金森病的转化。安。神经学。86, 357–367 (2019).

100.   Kidziński，Ł.，Delp，S.&Schwartz，M.。使用深度神经网络的儿童自动实时步态事件检测。《公共科学图书馆·综合》14，e0211466（2019年）。

101.   Díaz，S.，Stephenson，J.B.和Labrador，M.A.可穿戴传感器技术在步态、平衡和运动范围分析中的应用。申请。科学院。10, 234 (2020).

102.   Juen，J.，Cheng，Q.，Prieto Centurion，V.，Krishnan，J.A.&Schatz，B.通过移动电话步态分析对慢性病进行健康监测。电话。J、 E Health 20，1035–1041（2014年）。

103.   Kononova，A.等人，《在老年人中使用可穿戴活动跟踪器：跟踪器认知、动机和行为改变维持阶段障碍的焦点小组研究》。JMIR Mhealth Uhealth 7，e9832（2019年）。

104.   Da Gama，A.，Fallavollita，P.，Teichrieb，V.，&Navab，N.《使用Kinect进行运动康复：系统综述》。《奥运会健康杂志》4，123–135（2015）。

105.   Cho，C.-W.，Chao，W.-H.，Lin，S.-H.&Chen，Y.-Y.帕金森病患者步态识别的基于视觉的分析系统。专家系统。申请。36, 7033–7039 (2009).

106.   Seifert，A.，Zoubir，A.M.&Amin，M.G.使用室内多普勒雷达检测步态不对称。进行中。IEEE雷达会议1–6（IEEE，2019）。

107.   Altaf，M.U.B.，Butko，T.，Juang，B.H.&Juang，B.-H.声学步态：脚步声步态分析。IEEE传输。生物群落。2001年至2011年（2015年），第62版。

108.   Galna，B.等人，微软Kinect传感器测量帕金森病患者运动的准确性。步态姿态391062-1068（2014）。

109.   Jaume-i-Capó，A.，Martínez-Bueso，P.，Moyá-Alcover，B.&Varona，J.《改善脑瘫患者平衡疗法的互动康复系统》。

*IEEE传输。神经系统。康复。发动机*. 22, 419–427 (2014).

110.   Tinetti，M.E.，Williams，T.F.&Mayewski，R.《基于慢性残疾人数的老年患者跌倒风险指数》。是。J、 地中海。80, 429–434 (1986).

111.   Wang，C.等人。基于可穿戴惯性和麦克风传感器的多模态步态分析。进行中。IEEE智能世界1–8（2017）。

112.   心理健康美国。美国心理健康-2018年成人数据https://www.mhanational.org/issues/mental-健康-美国-成人-数据-2018（2018）。

113.   Wittchen，H.U.等人，《2010年欧洲精神障碍和其他大脑障碍的规模和负担》。欧元。神经精神药理学。21, 655–679 (2011).

114.   斯诺登，L.R.《心理健康评估和干预中的偏见：理论与证据》。

*是。J、 公共卫生*93, 239–243 (2003).

115.   Shatte，A.B.R.，Hutchinson，D.M.和Teague，S.J.心理健康中的机器学习：

方法和应用的范围审查。心理学。地中海。49, 1426–1448 (2019).

116.   Chakraborty，D.等。从RGB+D运动信号评估和预测精神分裂症阴性症状。进行中。第19届多媒体信号处理国际研讨会1–6（2017）。

117.   Pestian，J.P.等人，一项使用自然语言处理来检查急诊室自杀青少年语言的对照试验。自杀威胁生命。比哈夫。46, 154–159 (2016).

118.   Lutz，W.，Leon，S.C.，Martinovich，Z.，Lyons，J.S.&Stiles，W.B.《门诊心理治疗中的治疗师效应：三水平生长曲线法》。J、 康纳斯。心理学。54, 32–39 (2007).

119.   Miner，A.S.等人。评估心理治疗中自动语音识别的准确性。NPJ数字。地中海。3, 82 (2020).

120.   Xiao，B.，Imel，Z.E.，Georgiou，P.G.，Atkins，D.C.和Narayanan，S.S.“评估我的治疗师”：通过言语和语言处理自动检测药物和酒精咨询中的移情。PLoS ONE 10，e0143055（2015年）。

121.   Ewbank，M.P.等。使用深度学习量化心理治疗内容与临床结果之间的关联。《美国医学会精神病学杂志》77，35–43（2020）。

122.   Sadeghian，A.，Alahi，A.&Savarese，S.《追踪不可追踪的线索：学习追踪具有长期依赖性的多重线索》。进行中。计算机视觉和模式识别会议300–311（IEEE，2017）。

123.   刘，G。等。不规则孔洞的部分卷积图像修复。进行中。

*第15届欧洲计算机视觉会议*89–105（施普林格，2018）。

124.   Marafioti，A.，Perraudin，N.，Holighaus，N.&Majdak，P.用于音频修复的上下文编码器。IEEE/ACM传输。语音语言处理。27, 2362–2372 (2019).

125.   Chen，Y.，Tian，Y.&He，M.单目人体姿势估计：基于深度学习方法的调查。计算机。维斯。图像理解。192, 102897 (2020).

126.   Krishna，R.等人，《视觉基因组：使用众包密集图像注释连接语言和视觉》。国际计算机杂志。维斯。123, 32–73 (2017).

127.   Johnson，J.等人。使用场景图的图像检索。进行中。计算机视觉和模式识别会议3668–3678（IEEE，2015）。

128.   施，J.，张，H.&李，J.场景图的可解释和显式视觉推理。

进行中。计算机视觉和模式识别会议8368–8376（IEEE/CVF，2019）。

129.   哈兰卡，J.D.在一个学术医疗中心的大数据早期经验。卫生事务处。33, 1132–1138 (2014).

130.   Verbraeken，J.等人，《分布式机器学习综述》。ACM计算机。外科手术。53, 30 (2020).

131.   You，Y.等人，《深度学习的大批量优化：76分钟内训练BERT》。进行中。第八届学习表现国际会议1–38（2020）。

132.   基塔耶夫，N.，凯撒，Ł。&列夫斯卡娅，A.改革者：高效的变压器。进行中。

*第八届学习表征国际会议*1–12 (2020).

133.   Heilbron，F.，Niebles，J.&Ghanem，B.《快速时间活动建议》，用于有效检测未剪辑视频中的人类行为。进行中。计算机视觉和模式识别会议1914-1923（IEEE，2016）。

134.   Zhu，Y.，Lan，Z.，Newsam，S.&Hauptmann，A.用于动作识别的隐藏双流卷积网络。进行中。第14届亚洲计算机视觉会议363–378（Springer，2019）。

135.   Han，S.，Mao，H.&Dally，W.J.深度压缩：使用剪枝、训练量化和哈夫曼编码压缩深度神经网络。进行中。第四届学习表现国际会议1–14（2016）。

**介绍了一种对神经网络模型进行压缩的方法，降低了模型的计算量和存储量**.

136.   Micikeviciuds，P.等人，混合精度训练。进行中。第六届学习表现国际会议1–12（2018）。

137.   Yu，G.和Yuan，J.人类行为检测和搜索的快速行动建议。进行中。计算机视觉和模式识别会议1302–1311（IEEE，2015）。

138.   邹，J.&席宾格，L.艾可能是性别歧视者和种族主义者-是时候公平了。《自然》559324–326（2018）。

139.   Pleiss，G.，Raghavan，M.，Wu，F.，Kleinberg，J.&Weinberger，K.Q.《公平与校准》。副词。神经基础过程。系统。30, 5680–5689 (2017).

140.   Neyshabur，B.，Bhojanalli，S.，McAllester，D.&Srebro，N.。探索深度学习中的泛化。副词。神经基础过程。系统。30, 5947–5956 (2017).

141.   Howard，J.和Ruder，S.文本分类的通用语言模型微调。进行中。

*计算语言学协会第56届年会*328–339 (2018).

142.   潘世杰、杨问：关于迁移学习的调查。IEEE传输。我知道。数据工程22，1345–1359（2010）。

143.   Patel，V.M.，Gopalan，R.，Li，R.&Chellappa，R.。视觉域适应：最新进展综述。IEEE信号处理。杂志32、53–69（2015年）。

144.   王勇，郭杰，倪琳，姚琦。从几个例子概括：关于少数镜头学习的调查。ACM计算机。外科手术。53, 63 (2020).

145.   Jobin，A.，Ienca，M.&Vayena，E.《人工智能伦理指南的全球景观》。纳特。机器。内尔。1, 389–399 (2019).

146.   Li，C.，Lubecke，V.M.，Boric Lubecke，O.&Lin，J.非接触式医疗监测多普勒雷达传感器的最新进展综述。IEEE传输。微克。《理论技术》612046–2060（2013）。

147.   Rockhold，F.，Nisen，P.&Freeman，A.《十字路口的数据共享》。N、 英语。J、 地中海。375, 1115–1117 (2016).

148.   Wiens，J.等人，《无害：负责医疗机器学习的路线图》。纳特。地中海。25, 1337–1340 (2019).

149.   El Emam，K.，Jonker，E.，Arbuckle，L.&Malin，B.对健康数据再鉴定攻击的系统性回顾。《公共科学图书馆·综合》第6期，e28071（2011年）。

150.   Nasrollahi，K.&Moeslund，T.《超分辨率：综合调查》。机器。维斯。申请。25, 1423–1468 (2014).

151.   布鲁斯特，T。一个业余说唱乐队成员是如何窃取几乎追踪每个美国人的监控技术的。《福布斯》杂志（2018年10月12日）。

152.   卡特勒，J.E.患者如何从他们的医疗数据中赚钱？彭博社法（2019年1月29日）。

153.   Cahan，E.M.，Hernandez-Boussard，T.，Thadaney-Israni，S.&Rubin，D.L.将数据放在大数据解决个性化医疗的算法之前。NPJ数字。地中海。2, 78 (2019).

154.   Rajkomar，A.，Hardt，M.，Howell，M.D.，Corrado，G.&Chin，M.H.。确保机器学习的公平性以促进卫生公平。安。实习生。地中海。169, 866–872 (2018).

155.   Char，D.S.，Shah，N.H.和Magnus，D.在医疗保健中实施机器学习-解决道德挑战。N、 英语。J、 地中海。378, 981–983 (2018).

156.   Buolamwini，J.&Gebru，T.《性别阴影：商业性别分类中的交叉准确性差异》。进行中。第一届公平、问责制和透明度会议77–91（2018）。

157.   Chen，I.Y.，Szolovits，P.&Ghassemi，M.人工智能能帮助减少一般医疗和精神卫生保健方面的差异吗？AMA J.Ethics 21，E167–E179（2019年）。

158.   先证者：一种评估偏倚风险和预测模型研究适用性的工具。安。实习生。地中海。170, 51–58 (2019).

159.   Murdoch，W.J.，Singh，C.，Kumbier，K.，Abbasi Asl，R.&Yu，B.可解释机器学习的定义、方法和应用。过程。国家科学院。科学院。美国116，22071–22080（2019）。

**本文提出了一个通过预测准确性、描述准确性和相关性来评价模型可解释性的框架**.

160.   他，J。等。人工智能技术在医学中的实际应用。纳特。地中海。25, 30–36 (2019).

161.   Collins，G.S.，Reitsma，J.B.，Altman，D.G.&Moons，K.G.M.个体预后或诊断多变量预测模型（TRIPOD）的透明报告：TRIPOD声明。安。实习生。地中海。162, 55–63 (2015).

162.   Mitchell，M.等人，《模型报告模型卡》。进行中。第二届公平、问责制和透明度会议220–229（2019）。

163.   Thomas，R.等人，《协商民主与癌症筛查同意：社区陪审团对男性参与PSA筛查的知识和意愿影响的随机对照试验》。BMJ公开赛4，e005691（2014）。

164.   公共卫生实践不是研究。是。J。

*公共卫生*104, 596–602 (2014).

165.   Gerke，S.，Yeung，S.&Cohen，I.G.《医院环境智能的伦理和法律问题》。J、 是的。地中海。联合会第323、601–602号决议（2020年）。

166.   Kim，J.W.，Jang，B.&Yoo，H.个人健康数据流的隐私保护聚合。《公共科学图书馆·综合》第13期，第e0207639页（2018年）。

167.   van der Maaten，L.，Postma，E.&van den Herik，J.《降维：比较研究》。J、 马赫。学习。第10、13号决议（2009年）。

168.   Kocabas，M.，Athanasiou，N.&Black，M.J.VIBE：人体姿势和形状估计的视频推断。进行中。计算机视觉和模式识别会议5253–5263（IEEE/CVF，2020）。

169.   McMahan，H.B.，Moore，E.，Ramage，D.，Hampson，S.&Arcas，B.A。

从分散的数据中有效地学习深层网络。进行中。

*第20届人工智能与统计国际会议*1273–1282（PMLR，2017年）。

**本文提出了联邦学习，一种在数据分布在多个客户端设备上时训练共享模型的方法**.

170.   Gentry，C.使用理想格的完全同态加密。第41届计算机理论研讨会169–178（ACM，2009）。

**本文提出了第一个支持加密数据加法和乘法的完全同态加密方案**.

171.   麦考伊，S.T.登上美国海军舒适号（美国海军，2003年）。

**致谢**感谢A.Kaushal、D.C.Magnus、G.Burke、K.Schulman和M.Hutson对本文的评论。我们还感谢多年来我们的临床合作者，包括A.S.Miner、A.Singh、B.Campbell、D.F.Amanatullah、F.R.Salipur、H.Rubin、J.Jopling、K.Deru、N.L.Downing、R.Nazerali、T.Platchek和W.Beninati，以及多年来我们的技术合作者，包括A.Alahi、A.Rege、B.Liu、B.Peng、D.Zhao、E.Chou、E.Adeli，比安科尼、普西奥、蔡浩、比亚尔、谢剑廷、郭明义、梅赫拉、杨圣贤、罗志强。A、 H.的研究生工作得到了美国海军研究办公室（批准号：N00014-16-1-2127）和斯坦福人工智能研究所的部分支持。

**作者贡献**A、 H，A.M.和L.F-F.概念化了论文及其结构。A、 H.和L.F.-F.写了这篇论文。A、 H.创造人物。A、 提供了大量的补充和编辑。所有作者都对论文的多个部分以及最终的风格和总体内容作出了贡献。

**相互竞争的利益**A、 M.在Prealize Health有经济利益。五十、 F.-F.和A.M.在Dawnlight Technologies有经济利益。A、 H.声明没有相互竞争的利益。

**附加信息**

**补充资料**可在https://doi.org/10.1038/s41586-0202669-y。

**信函和材料要求**应致L.F.-F.同行评审信息自然感谢Andrew Beam、Eric Topol和其他匿名评审人对本工作同行评审的贡献。

**再版和权限信息发布者说明**可在http://www.nature.com/reprints。斯普林格自然保持中立，就管辖权的要求，在出版的地图和机构的联系。

©施普林格自然有限公司2020

[[1]](" \l "_ftnref1" \o ")斯坦福大学计算机科学系，斯坦福，加利福尼亚，美国。斯坦福大学医学院临床卓越研究中心，斯坦福，加利福尼亚，美国。斯坦福人为中心的人工智能研究所，斯坦福大学，斯坦福，加利福尼亚，美国。✉电子邮件：feili@stanford.edu23