🖊 有道翻译

研究ARTICLES▼

经济学

结合卫星图像和机器学习来预

NealJean, 1,2*MarshallBurke, 3,4,5* † Michael Xie, 1W. Matthew Davis, 4David B. Lobell, 3,4Stefano Ermon1

在发展中国家、有关经济生计的可靠数据仍然稀缺、阻碍了研究这些成果和设计改善这些成果的政策的努 力。本文展示了一种准确、廉价和可扩展的方法,用于从高分辨率卫星图像中估计消费支出和资产财富。 利用来自五个非洲国家(尼日利亚、坦桑尼亚、乌干达、马拉维和卢旺达)的调查和卫星数据,我们展示了 如何训练卷积 经网络来识别图像特征,这些特征可以解释高达75%的地方层面经济结果差异。我们的方 角的数据,可以改变跟踪和定位发展中国家贫困的努力。它还展示了机器学习技术在训练 限的情况下可以应用的多么强大,表明了在许多科学领域的广泛潜在应用。

人口经济特征的准确测量 对研究和政策都有至关重要的影响。这些测 量结果决定了各国政府如何分配稀缺资源, 并为了解和跟踪改善人类生计进展的全球努 力奠定了基础。尽管近年来发展中国家可获 得的经济数据的数量和质量都有所提高,但 对大多数发展中国家来说, 经济发展关键指 标的数据仍然缺乏(1)。这一数据差距阻碍了 识别和了解这些结果的差异以及将干预有效 地定向到最需要的领域的努力(2,3)。

非洲大陆的数据差距尤其令人制约。根 据世界银行的数据,在2000年至2010年期间, 59个非洲国家中有39个国家进行了少于两项 可以构建具有国家代表性的贫困衡量指标的 调查。在这些国家中,有14个国家在此期间 没有进行此类调查(4)(图1A), 所进行调查的 大多数数据都不在公共领域。人口与健康调 查(DHS)的覆盖面同样有限,人口与健康调 查是大多数发展中国家人口水平健康统计数 据的主要来源,也是家庭资产(一种常见的 财富衡量标准)的国际可比数据的主要来源 (图1B)。在同样的11年期间,59个国家中有 20个是贫困国家

尝试没有进行基于DHS资产的调查, 另外19 个只有一次。这些缺点促使人们呼吁进行 "数据革命",以大幅扩大在非洲和其他地 方的数据收集工作(1)。但通过更频繁的家庭 调查来弥补这些数据差距可能不仅成本过 高——在每个国家15年期间衡量联合国可持 续发展目标的每个目标可能要花费数千亿美 元(5)——而且在制度上也存在困难,因为一 些国家的政府认为, 记录在案的黯淡表现没 有什么好处(2,6)。

考虑到扩大传统数据收集工作的困难, 衡量这些结果的另一种途径可能是使用被动 收集数据的新来源,如来自社交媒体、移动 电话网络或卫星的数据。最近流行的一种方 法利用夜间光度的卫星图像("夜灯")来估 计经济活动(7-10)。虽然这一特殊技术已显 示出改善现有国家一级经济生产统计数据的 希望(7,10), 但它似乎不太能区分人口生活 在国际贫困线附近或以下的地区(人均每天1. 90美元)的经济活动差异。在这些贫困地区, 光度水平通常也很低,变化不大(图1、C到F 和图S1), 这使得夜间灯光对研究和跟踪非 常贫困的人的生计可能不太有用。其他最近 使用手机数据估计贫困的方法(11,12)显示出 了希望, 但由于它们依赖于不同的专有数据 集,可能难以在国家之间扩展。

本文展示了一种新颖的机器学习方法, 用于从高分辨率日间卫星图像中提取社会经 济数据。然后,我们在五个非洲国家验证了 这种方法, 这些国家最近的地理参考当地水 平数据

可以获得经济成果。与现有方法相比,我们 的方法仅使用公共领域可用的数据就可以产 生细粒度的贫困和财富估计。

转移学习

高分辨率卫星图像在全球范围内的可用性日 益增加,其中包含了可以与经济活动相关联 的丰富景观特征信息。不幸的是, 这些数据 是高度非结构化的, 因此即使进行密集的人 工分析, 也很难从大规模中提取有意义的见 解。最近深度学习技术在大规模图像数据集 上的应用已经导致了基础计算机视觉任务 (如目标检测和分类)的显著改进, 但这些技 术通常在标记训练数据丰富的监督学习机制 中最有效(13)。然而,在我们的设置中,带 标签的数据是稀缺的。即使在详细的家庭调 查确实存在的情况下(图1,A和B),单个调查 通常只包含数百个地点的信息,产生的数据 集比通常在深度学习应用中使用的数据集小 许多个数量级。因此,尽管卷积神经网络等 深度学习模型原则上可以通过训练直接估计 卫星图像的经济成果, 但这些成果的训练数 据的稀缺使得这些技术的应用具有挑战性。

我们通过一种多步骤的"迁移学习"(14)方 法克服了这一挑战(参见补充材料第1节),即 使用一个嘈杂但容易获得的贫困的代理来训 练深度学习模型(15)。然后,该模型被用于 估计"集群"水平上的平均家庭支出或平均 家庭财富(大致相当于农村地区的村庄或城 市地区的小区), 这是我们使用的公共领域 调查中纬度和经度数据可用的最低地理聚集 水平(见补充材料1.4)。在可以获得的情况下, 家庭支出是贫穷国家计算国家贫困统计数据 的标准依据, 我们使用的支出数据来自世界 银行的生活水平测量研究(LSMS)调查。为 了衡量财富, 我们使用了从美国国土安全部 (DHS)中提取的资产指数,作为对有关资产 所有权的多个问题的调查答复的第一主成分。 尽管资产指数不能直接用于构建贫困的基准 测量, 但基于资产的测量被认为可以更好地 捕捉家庭的长期经济状况(16,17), 还有一个 额外的优势, 即许多枚举的资产可以直接被 测量者观察到, 因此测量误差相对较小。

为了估计结果, 我们的迁移学习流程包 括三个主要步骤。首先, 我们从卷积神经网 络(CNN)模型开始,该模型已经在imagenet 上进行了预训练, imagenet是由标记图像组 成的大型图像分类数据集

Department of Computer Science, Stanford University, Stanford, CA, USA. Department of Electrical Engineering, Stanford University, Stanford, CA, USA. Department of Earth System Science, Stanford University, Stanford, CA, USA.

⁴Center on Food Security and the Environment, Stanford University, Stanford, CA, USA. ⁵National Bureau of Economic Research, Boston, MA, USA.

^{*}These authors contributed equally to this work. †Corresponding author. Email: mburke@stanford.edu

来自1000个不同类别(18个)。在学习正确分类每个图像(例如, "仓鼠"与"黄鼠狼")时,模型学习识别许多视觉任务共同的边缘和角点等低级图像特征(19)。

接下来, 我们在从这个图像分类任务中获 得的知识的基础上,对CNN在一个新的任务上 进行微调, 训练它预测与输入的白天卫星图像 对应的夜间光线强度。在这里, 我们使用"预 测"一词来表示对某些未直接观察到的属性的 估计, 而不是其通常的含义——推断有关未来 的某些东西。夜灯是经济活动的一种嘈杂但在 全球范围内一致的——而且在全球范围内可用 的——代表。在第二步中,模型学习将高维输 入的白天卫星图像"总结"为一组较低维的图 像特征,这些特征可以预测夜间灯光的变化(见 图2)。训练后的CNN可以被视为一个特征提取 器,它已经学习了从每个输入图像到简洁特征 向量表示的非线性映射(补充材料1.1)。日间图 像(此处取自谷歌静态地图API)和夜灯(20)都可 以以相对较高的分辨率获得整个全球地表、提 供了一个非常大的标记训练数据集。

最后,我们使用调查数据中的平均聚类水平值以及CNN从日间图像中提取的相应图像特征来训练山脊回归模型,该模型可以估计聚类水平的支出或资产。ridge模型中的正则化可以防止过拟合,鉴于所提取特征的高维性和相对较小的调查数据集,这是一个潜在的挑战。直观地说,我们期望解释夜间灯光变化的一些特征子集也可以预测经济结果。

在不完美的经济幸福指标上部分训练的模型——在这种情况下,在上述第二步训练中使用的夜间灯光——如何改善直接使用该指标作为幸福指标的评估器?虽然夜间照明在较低的支出水平下变化不大(图1,C至F),但调查数据表明,白天卫星图像中可见的其他特征,如屋顶材料和到城市地区的距离,大致随支出呈线性变化(图S2),因此可以更好地捕捉到较贫穷集群之间的变化。因为夜灯和这些特征都显示了在较高收入水平下的变化,对夜灯进行训练可以帮助CNN学习提取这样的特征,更有能力捕捉整个消费分布的变化。

夜灯也很难区分贫穷、人口密集的地区和富裕、人口稀疏的地区,这是不使用夜灯来估计人均消费的额外动机。 我们的方法不依赖于夜灯能够做出这种区分,而是只使用夜灯作为中间标签来学习与经济福祉相关的图像特征。我们分析的最后一步,在这一步中我们训练模型直接

从白天估算当地人均产出 image features, does not rely on nightlights.

Visualization of the extracted image features suggests that the model learns to identify some livelihood-relevant characteristics of the landscape (Fig. 2). The model is clearly able to discern semantically meaningful features such as urban areas, roads, bodies of water, and agricultural areas, even though there is no direct supervision—that is, the model is told neither to look for such features, nor that they could be correlated with economic outcomes of interest. It learns on its own that these features are useful for estimating

夜间灯光强度。这与

to existing efforts to extract features from satellite imagery, which have relied heavily on humanannotated data (21).

Results

Our transfer learning model is strongly predictive of both average household consumption expenditure and asset wealth as measured at the cluster level across multiple African countries. Crossvalidated predictions based on models trained separately for each country explain 37 to 55% of the variation in average household consumption

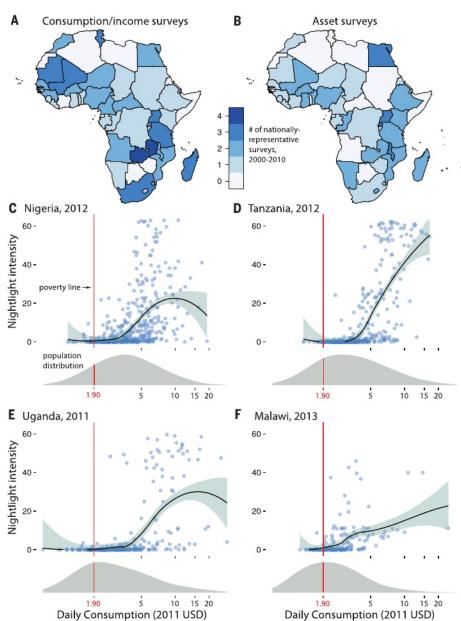


图1所示。贫困数据差距。(A) 2000年至2010年期间每个非洲国家进行的具有全国代表性的消费调查次数。(B)与(A)相同,用于衡量资产的人口与健康调查。(C至F)根据住户调查,四个非洲国家的人均消费支出(以美元计算)与群集一级夜间灯光强度之间的关系。消费分布中每个点上具有全国代表性的家庭所占份额在每个面板下方以灰色表示。垂直红线表示官方的国际极端贫困线(每人每天1.90美元),黑线用浅蓝色对应的95%置信区间拟合数据。

SCIENCE sciencemag.org 19 AUGUST 2016 · VOL 353 ISSUE 6301 791

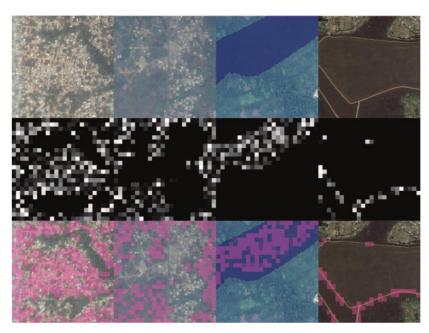


图2所示。特征的可视化。按列:用于提取特征的卷积神经网络模型中,四个不同的卷积滤波器(从左到右分别识别城市区域、非城市区域、水和道路对应的特征)。每个滤波器"高亮"图片中激活它的部分,用粉色表示。按行:原始的白天卫星图像来自谷歌静态图,滤镜激活图,以及叠加在原始图像上的激活图

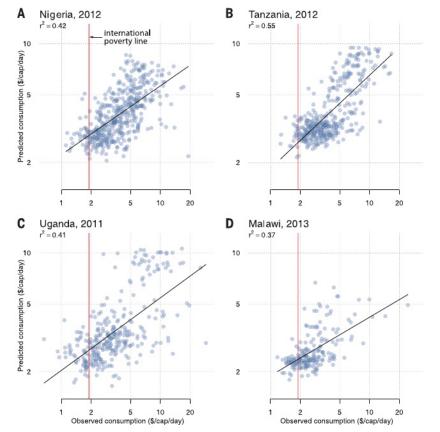


图3所示。从迁移学习方法预测的集群级消费(y轴)与调查测量的消费(x轴)比较。结果显示了尼日利亚(A)、坦桑尼亚(B)、乌干达(C)和马拉维(D)。每个面板中的预测和报告的r·值来自五折交叉验证。黑线为最佳拟合线,红线为每人每天1.90美元的国际贫困线。两个坐标轴均以对数尺度表示。国家按人口数量排序。

4个国家的最新调查数据(图3),以及5个国家的平均家庭资产财富变化的55-75%(图S3)。在所有国家的汇集消费或资产观察上训练的模型(以下简称"汇集模型")表现类似,交叉验证的预测解释了这些结果的总体变化的44-59%(图S4)。

尽管缺乏日间图像的时间标签(即每张 图像的确切日期未知), 以及对簇的位置不 完全了解, 但仍然实现了这种高总体预测 能力, 因为数据收集机构为保护受访者的 隐私,在簇坐标中添加了多达10公里的随 机噪声。资产的预测能力几乎一致高于消 费的预测能力,这可能反映了资产调查中 可用的样本量较大;资产指数被认为可以更 好地反映家庭的长期经济状况(16,17)(这可 能与随着时间缓慢变化的地貌特征更好地 相关):和/或指数中的某些资产(如屋顶类型) 在提取的特征中被直接识别的可能性(见补 充材料2.1)。我们通过从乌干达lsm中可用 的变量构建我们自己的资产指数, 并将该 指数的预测性能与同一调查中测量的消费 性能进行比较, 来研究这些潜在的解释。 我们发现,被测量结果的差异,而不是调 查设计或直接识别keyassetsindaytimeimager y的差异,可能解释这些性能差异(见补充材 料2.1和图S5)。最后, 我们的模型在卢旺达 的资产估计性能超过了最近的一项研究中 使用手机数据估计相同结果的性能(11)(该 研究中的集群级r²= 0.62, 我们的研究中r²= 0.75;R²是决定系数),同样具有额外优势的 是, 我们的预测可以完全从公开可用的数 据中构建, 在跨国家扩展时无需获取和评 估专有数据集。

为了测试我们的迁移学习模型在直接使用夜灯来估计生计时是否有所改进,我们分别对每个国家和联合模型进行了100次10折交叉验证试验,每次都将我们的迁移学习模型的预测能力与单独使用夜灯的预测能力进行比较。为了了解在消费分布的不同子集上的相对表现,试验是单独运行的,将集群的样本限制在平均消费低于消费分布的每个五分位数的集群上。对资产重复执行相同的程序。

尽管在夜间灯光上进行了部分训练,但平均而言,我们的模型比单独的夜间灯光对消费和资产变化的预测能力要高得多。对于支出,我们的模型在消费分布的几乎所有点上都优于夜灯,无论是对于集合模型还是独立运行的国家(图4A和图S6)。在集合模式下,for

792 19 AUGUST 2016 · VOL 353 ISSUE 6301 sciencemag.org SCIENCE

在国际贫困线以下的集群中, 我们的模型在 81.3%的试验中优于夜灯, r2的平均增幅为0. 04。对于低于两倍贫困线的集群,我们的模 型在98.5%的试验中优于夜灯,r2的平均增幅 为0.10、解释力增加了81.2%。对于低于贫困 线三倍的集群, 我们的模型在99.5%的试验 中优于夜灯, r²平均增加0.12, 对应解释力增 加54.2%。单个国家的结果是相似的,我们 的模型的预测能力在消费分布的几乎所有部 分都超过了所有国家的夜灯(图S6)。我们的 模型相对于nightlights的相对表现甚至比消费 更好(图4B), 特别是对于平均资产水平较低 的集群。在夜灯中使用超出平均亮度的更多 信息会导致夜灯性能的一些提高, 但这种对 夜灯的改进使用仍然被我们的模型所超越(参 见补充材料2.2和图S7)。

我们还研究了我们的方法是否比其他更简单的方法更好,可以从日间图像中提取信息,并使用可用的调查数据预测经济结果。我们发现,我们的CNN特征提取器的表现远远优于常见的通用图像特征,如颜色直方图和方向梯度直方图(见补充材料

2.3和图S8)。我们的方法在最近的调查中也 表现得与使用过去调查的数据来预测结果的 直观方法一样好或更好(见补充材料)

2.4和表S2)。

为了进一步量化结果的统计显著性,进行了一项实验,将日间图像随机重新分配到调查地点,并在这些不正确的图像上重新训练模型(见补充材料1.7)。我们在每个国家和集合模型中重复这个实验1000次,然后将白天图像分配到正确位置(如图3所示)时的预测能力与从1000个安慰剂试验中获得的r²值的分布进行比较。如图4,C和D所示,无论是消费还是资产,使用"正确"的日间图像获得的r²值都远高于从重新洗牌的图像中获得的任何r²值,这表明我们的模型的预测性能水平不太可能是偶然产生的。

最后,利用我们对多个国家的消费和资产的基于调查的测量,我们研究了使用一个国家的数据和卫星图像特征训练的模型在多大程度上可以估计其他国家的生计。研究某一特定模型是否能跨国界推广,有助于了解在没有调查数据的地区是否能仅从图像中做出准确的预测(鉴于许多非洲国家现有调查数据不足,这是一个重要的实际问题),也有助于了解各国生计决定因素的共性。

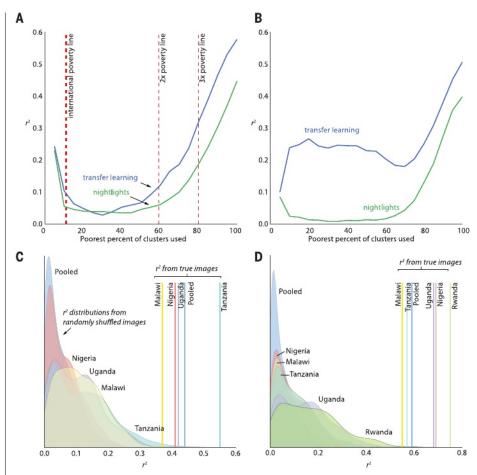


图4所示。模型性能评估。(A)使用对四个LSMS国家的汇总观察,估算消耗的迁移学习模型相对于夜间灯光的表现。 为了增加可用集群的百分比,分别进行了试验(例如,x轴值为40表示包括了消耗低于40百分位数的所有集群)。垂直 红线表示国际资困线的各种倍数。利用主成分分析将图像特征降至100维。(B)与(A)相同,但针对资产。(C)将在每 个国家正确分配的图像上训练的模型+(垂直线)与在随机打乱的图像上训练模型(每个国家1000次试验)的试验中获得 的+值的分布进行比较。(D)与(C)相同,但针对资产。在所有面板中报告交叉验证的+值。

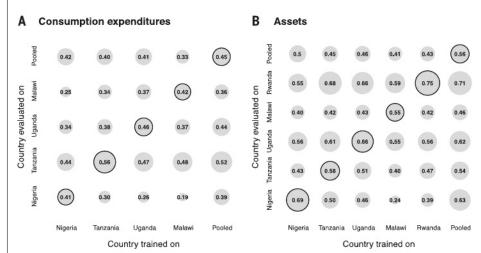


图5所示。跨界模型泛化。(A)在一个国家训练并在其他国家应用的模型的消费预测交叉验证产值。x轴上的国家表示模型训练的地方,y轴上的国家表示模型评估的地方。报告的产值在100倍以上被平均(10次试验,每次10倍)。(B)与(A)相同,但针对资产。

SCIENCE sciencemag.org 19 AUGUST 2016 · VOL 353 ISSUE 6301 793

我们发现,无论是消费还是资产,在国内训练的模型都一致优于在国外训练的模型(图5),这与预期一致。但我们也发现,模型似乎在跨越国界时"旅行得很好",国外的预测往往接近国内预测的准确性。在所有四项消费调查或所有五项资产调查中训练的汇集模型,对这两种结果的预测能力几乎接近几乎所有国家的国内模型的预测能力。这些结果表明,至少就我们的样本国家而言,生计的共同决定因素在图像中得到了揭示,在未能观察到调查结果的国家,可以利用这些共同点以合理的准确性估计消费和资产结果。

讨论

我们的方法表明,现有的高分辨率日间卫星图像可以用来对五个非洲国家的经济福祉的空间分布做出相当准确的预测。尽管白天图像的时间和簇在训练数据中的位置都不精确,但我们的模型表现良好,这些维度中的任何一个更精确的数据都可能进一步提高模型的性能。

值得注意的是,我们表明,<mark>当在我们的一个样本国家训练的模型被用于估计另一个国家的消费或资产时,我们的模型的预测能力只会适度下降。尽</mark>管各国的经济和政治制度存在差异,但模型衍生的特征似乎识别出了不同环境下生计决定因素的基本共性,这表明我们的方法可以用来填补许多非洲国家因调查覆盖率低而造成的巨大数据空白。与最近其他依赖专有商业数据集的方法相比,我们的方法只使用公开可用的数据,因此在各国之间扩展是直接的,几乎没有成本。

虽然所提出模型在估计集群水平的经济福祉方面优于其他被动收集的数据来源(如手机数据、夜间灯光),但目前无法评估其区分集群内差异的能力,因为公共领域调查数据为给定集群中的所有家庭分配了相同的坐标,以保护受访者隐私。原则上,我们的模型可以在任何分辨率下对可用的日间卫星图像进行预测,尽管在更精细的尺度上进行预测可能会有更大的噪声。新的真实数据来源,无论是来自更分散的调查还是新颖的众包渠道,都可以在家庭层面对我们的模型进行评估。将我们提取的特征与其他被动收集的数据相结合,在这些数据可用的位置,也可以增加家庭和集群级别的预测能力。

考虑到白天图像的高分辨率时间序列的可 用性有限,我们还无法评估我们的迁移学习方 法预测特定位置随着时间的推移经济福祉变化 的能力。这

794

样的预测对研究人员和政策制定者都有帮助, 并且在不久的将来,随着越来越多的高分辨率 卫星图像可用,应该能够实现(22)。

我们使用丰富但嘈杂的代理的迁移学习策略表明,通常在数据丰富的环境中蓬勃发展的强大的机器学习工具,即使在感兴趣的关键结果的数据稀缺的情况下,也可以有效地使用。我们的方法可以在许多科学领域广泛应用,并可能立即有用,以低廉的成本产生有关国际社会感兴趣的其他社会经济结果的细粒度数据,例如为联合国可持续发展目标(5)提出的大量指标。

参考资料和说明

- 1. United Nations, "A World That Counts: Mobilising the Data Revolution for Sustainable Development" (2014).
- 2. S. Devarajan, Rev. Income Wealth 59, S9-S15 (2013).
- M. Jerven, Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It (Cornell Univ. Press, 2013).
- World Bank, PovcalNet online poverty analysis tool, http://iresearch.worldbank.org/povcalnet/ (2015).
- M. Jerven, "Benefits and costs of the data for development targets for the Post-2015 Development Agenda," Data for Development Assessment Paper Working Paper, (Copenhagen Consensus Center, Copenhagen, 2014).
- 6. J. Sandefur, A. Glassman, J. Dev. Stud. 51, 116-132 (2015).
- J. V. Henderson, A. Storeygard, D. N. Weil, Am. Econ. Rev. 102, 994–1028 (2012).
- X. Chen, W. D. Nordhaus, Proc. Natl. Acad. Sci. U.S.A. 108, 8589–8594 (2011).
- 9. S. Michalopoulos, E. Papaioannou, Q. J. Econ. 129, 151–213 (2013).
- M. Pinkovskiy, X. Sala-i-Martin, Q. J. Econ. 131, 579–631 (2016).

- 11. J. Blumenstock, G. Cadamuro, R. On, Science 350, 1073–1076
- L. Hong, E. Frias-Martinez, V. Frias-Martinez, "Topic models to infer socioeconomic maps," AAAI Conference on Artificial Intelligence (2016).
- 13. Y. LeCun, Y. Bengio, G. Hinton, Nature 521, 436-444 (2015).
- S. J. Pan, Q. Yang, IEEE Trans. Knowl. Data Eng. 22, 1345– 1359 (2010).
- M. Xie, N. Jean, M. Burke, D. Lobell, S. Ermon, "Transfer learning from deep features for remote sensing and poverty mapping." AAAI Conference on Artificial Intelligence (2016).
- 16. D. Filmer, L. H. Pritchett, Demography 38, 115-132 (2001).
- 17. D. E. Sahn, D. Stifel, Rev. Income Wealth 49, 463-489 (2003).
- 18. O. Russakovsky et al., Int. J. Comput. Vis. 115, 211-252 (2014).
- A. Krizhevsky, I. Sutskever, G. E. Hinton, Adv. Neural Inf. Process. Syst. 25, 1097–1105 (2012).
- National Geophysical Data Center, Version 4 DMSP-OLS Nighttime Lights Time Series (2010).
- V. Mnih, G. E. Hinton, in 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5 to 11 September 2010 (Springer, 2010), pp. 210–223.

致地: Hand, Science 348, 172-177 (2015).

教花戲謝NVIDIA公司通过NVIDIA学术硬件补助金、斯坦福大学全球发展与贫困倡议以及威廉与玛丽学院的AidData项目的支持。新泽西州感谢国防科学与工程研究生奖学金项目的支持。S.E.通过分包合同72954-10597获得NSF 1522054赠款的部分支持。我们声明没有利益冲突。复制这些结果所需的所有数据和代码都可以在http://purl.stanford.edu/cz134jc5378上找到。

www.sciencemag.org/content/353/6301/790/suppl/DC1 材料与方法

无花果。S1至S22

表S1到S3

引用(汽车)

2016年3月30日;2016年7月6日录取 10.1126/science.aaf7894

统计物理学

孤立多体系统中通过纠缠的 量子热化

Adam M. Kaufman, M. Eric Tai, Alexander Lukin, Matthew Rispoli, Robert Schittko, Philipp M. Preiss, Markus Greiner*

统计力学依赖于热平衡状态下系统的熵最大化。然而,在纯态中初始化的孤立量子多体系统在Schrödinger演化过程中保持纯态,从这个意义上说,它具有静态的、零熵。我们通过实验研究了量子态中统计力学的出现,并观察到量子。加速在促进这种出现方面的基本作用。进化中的量子系统的显微镜观察表明,全量子态保持纯净。加速上则发生在局域尺度上。我们直接测量了纠缠熵,它承担了热熵在热化中的作用。纠缠产生的局部熵验证了统计特理对局部可观测物的使用。我们的测量结果与本征态热化假说一致。

一个孤立的量子系统受到扰动时——例如,由于哈密顿量的突然变化(所谓的猝灭)——随后的动力学是由猝灭引起的本征态分布决定的(1)。在任何给定的时间,演化的量子态将具有

振幅依赖于由淬灭和哈密顿量的能量特征值 填充的特征态。然而,在许多情况下,

Department of Physics, Harvard University, Cambridge, MA 02138, USA.

19 AUGUST 2016 · VOL 353 ISSUE 6301 sciencemag.org SCIENCE

^{*}Corresponding author. Email: greiner@physics.harvard.edu



结合卫星图像和机器学习预测贫困

Neal Jean, Marshall Burke, Michael Xie, W. Matthew Davis, David B. Lobell和 Stefano Ermon(2016年8月18日)

《科学》353(6301), 790-794。(doi: 10.1126 / science.aaf7894)

编者总结

远程测量消费和财富

夜间照明是衡量经济财富的一个粗略指标,世界夜间地图显示,许多发展中国家的照明很少。Jean等人将夜间地图与高分辨率的日间卫星图像相结合(参见Blumenstock的视角)。通过一点机器学习的魔法,组合后的图像可以转换为家庭消费和资产的精确估计,而这两者在较贫穷的国家是很难测量的。此外,夜间和白天的数据是公开的,非专利的。

Science, 本期第790页;另见第753页

此副本仅供您个人使用, 非商业用途。

工具条 权限 访问本文在线版,获取个性化和文章工具:http://science.sciencemag.org/content/353/6301/790

获取转载本文相关信息:http://www.sciencemag.org/about/ permissions.dtl

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title Science is a registered trademark of AAAS.