

动态、高分辨率贫困 数据稀缺环境中的测量

卓正蒂莫西·
吴理查德·李·
大卫·纽豪斯·
塔利普·基利
奇·马歇尔·伯
克·斯特凡诺·
埃蒙·大卫·B
·洛贝尔



WORLD BANK GROUP

发展经济学

发展数据组

2025 年 2 月

Abstract

准确而全面的家庭生计测量对于监测减贫进展和社会援助计划的针对性至关重要。然而，传统数据收集的高成本历来使全面测量成为一个难题。本文评估了四种非洲国家详细家庭普查提取数据的替代基于卫星的深度学习方法，以加速大规模、细粒度和动态贫困测量的进步。结果表明，变压器架构解决了多个开放问题。

测量问题，通过提供对各国和城市家庭资产财富局部水平变化的最准确测量，以及家庭资产财富随时间的变化进行测量。人工限制数据可用性的实验展示了模型在数据有限的情况下实现高性能的能力。提出的方法证明了结合卫星图像、公开可获取的地理特征以及新的深度学习架构，可以在数据稀缺的环境中实现高度局部化和动态的贫困测量的潜力。

这篇论文是由发展数据组、发展经济学部门完成的。它是世界银行为提供研究的开放访问并为全球发展政策讨论做出贡献的一项更大努力的一部分。政策研究工作论文也在网上发布，网址为<http://www.worldbank.org/prwp>。作者可以联系dnewhouse@worldbank.org和tkilic@worldbank.org。

该政策研究工作论文系列发布正在撰写中的研究成果以促进关于发展方面的思想交流。问题。该系列的一个目标是迅速发布研究成果，即使展示内容尚未完全润色。论文保留了原始状态。作者的名字，并且应当据此引用。本文中的发现、解释和结论完全是作者个人观点。由作者提供。它们不一定代表国际复兴开发银行/世界银行的观点。其附属组织，或世界银行执行董事或他们所代表的政府的组织。

数据稀缺环境中的动态高分辨率贫困测量

Zhuo 郑, ^a 蒂莫西 Wu, ^a 理查德 Lee, ^{b,c} 大卫 纽豪斯, ^d Talip Kilic, ^d 马歇尔 伯克, ^c
^e Stefano Ermon, ^a and 大卫 B. Lobell ^{b,c}

^a 斯坦福大学计算机科学系, 斯坦福大学, 加利福尼亚州 94305, 美国

^b 斯坦福大学地球系统科学系, 斯坦福大学, 加利福尼亚州 94305, 美国

^c Center on Food Security and the Environment, Stanford University, Stanford, 94305, CA, USA

^d Development Economics Data Group, World Bank Group, Washington DC, 20433, DC, USA

^e Stanford University, Stanford, 94305, CA, USA

ARTICLE INFO

关键词 : Economicwell

- being 高分辨率贫困卫星

图像 Deep learning

JEL 代码: C45, I32

准确、最新且高度细化的经济福祉测量对于监测和实现国际减贫目标至关重要。这些目标包括联合国可持续发展目标1 (No Poverty), 该目标的原始截止日期为2030年, 以及无数其他国际和地区减贫目标。对于理解这些目标是否正在实现, 以及在进展滞后地区对减贫干预措施进行定位和评估而言, 家庭贫困和财富的细粒度估计至关重要。 [6]

官方对低收入和中等收入国家的贫困测量长期以来依赖于家庭调查, 这是一种不可或缺但耗时较长的生计测量工具。鉴于可靠调查测量所需的技术水平以及在进行全国或次区域代表性的生计调查时面临的重大后勤挑战, 此类调查往往在世界许多地区很少完成, 导致在许多时期和地区无法获得全面和及时的贫困及相关结果的衡量数据。 [6, 9] 与此同时, 调查数据通常基于旨在更大空间尺度上具有代表性的样本, 因此通常无法生成村庄或小区层面的可靠估计——而反贫困干预措施往往需要在这些层面进行针对性部署。因此, 迫切需要更多成本效益高且可扩展的替代方案来衡量生计状况, 以补充和扩大现有的基于家庭调查的努力。

在近年来, 公开可用的遥感数据日益丰富, 以及机器学习的最新进展已逐步改变了生计测量的格局, 逐渐从国家人口普查和相关家庭调查转向结合这些信息与卫星和其他传感器的数据。早期的研究使用的是粗略的、公开可获取的数据。

利用可用的卫星图像和/or移动电话数据, 结合早期的机器学习和深度学习架构, 展示这些新型信息源如何被用于支持大规模的财富和贫困测量。 [4, 17, 29]。subsequent 研究引入了进一步的改进, 这些研究利用了公开可用或专有的地理空间数据来提高基于卫星的财富测量精度 [8, 12, 2]。这些进步证实了利用卫星图像和机器学习可以提供一种准确、经济且可扩展的解决方案来估算

财富 [6, 21, 23]。

财富 [

我们汇集了一个大规模、多分辨率和多时间尺度的贫困数据集, 使用了国家普查数据或从国家统计局获取的提取数据, 并结合了来自不同分辨率多种公共和私营传感器的多光谱卫星影像。该数据集涵盖了四个非洲国家 (马拉维、莫桑比克、布基纳法索和马达加斯加) 超过120万户家庭的数据, 并且独特地包含两个马拉维城市内精确地理参考的测量数据, 以及同一地点随时间的重复测量数据——这是此前研究中所缺乏的特征。

我们利用这些数据做出了四项贡献, 相对于早期的研究工作。首先, 我们直接测试了一种新的深度学习模型——具体来说, 是视觉变换器 (Vision Transformers) ——与文献中常见的基于卷积神经网络 (CNNs) 的早期深度学习架构进行对比, 同时也与使用地理空间特征和表格机器学习方法 (如XGBoost) 进行预测的更简单的模型进行对比。具体而言, 我们设计了一个条件模块, 使我们的变压器模型能够处理多模态输入, 同时整合卫星影像和地理空间特征 (参见“方法”部分)。我们测试了使用不同分辨率的卫星影像数据构建的模型, 包括来自Landsat (每像素30米)、PlanetScope (每像素3米) 和SkySat (每像素0.5米) 传感器的数据。然后, 我们将这些更为复杂的方法和输入与依赖于更简单方法的模型进行了比较。

仅在一系列来源的预定义地理空间特征上。

这些比较非常重要，因为依赖公开数据的更简单的做法在大规模应用中可能既更容易实施也更经济，尤其是在公共组织希望广泛应用于其领域时。因此，理解不同模型架构和输入下的性能权衡对于理解如何规模化推广有前景的新测量方法至关重要。

其次，在我们的情境中，一个关键优势在于使用来自国家人口普查或提取的准确且高分辨率的数据进行模型训练和评估。与早期依赖于空间定位精度较低且家庭样本有限的公开可用的家庭调查数据的研究相比，我们的数据覆盖了给定位置上规模更大的家庭集，并且在某些情况下可以精确地理参考。与这些“金标准”数据进行比较，使我们能够理解模型预测误差是由于预测不准确还是真实地面数据测量中的噪声所致——这一点在早期发展中国家的研究中往往难以明确区分。⁶，²⁹ 此外，这使我们能够考虑广泛的样本大小范围，以评估先进机器学习方法产生准确估计所需的最小训练数据要求。为了量化训练样本大小对性能的重要性，我们在多个设置中广泛测试额外训练数据对模型性能的影响。

第三，我们高分辨率的人口普查数据使我们能够对如何利用卫星图像和其他地理空间数据来预测非洲城市区域内生计差异获得一种新的理解。这一能力在以往的研究中难以评估，因为训练数据样本有限且存在空间噪声。即使在小范围内，城市内部也表现出显著的空间生计差异性。通过使用马拉维两国城市全面且精确地理参照的人口普查数据，我们能够使用不同分辨率的卫星数据进行模型训练和测试，并发现这些模型在预测这些城市街道和街区层面的财富差异方面出乎意料地准确。

第四，通过人口普查和抽样数据，我们能够评估基于图像的模型是否能够准确预测财富随时间变化的情况。此前的努力再次受到地面数据的显著限制，这些数据未能在不同调查中反复采样相同的地点^[29] 因此，尚不清楚一种主要训练用于预测财富或消费在空间上变化的基于图像的模型是否能够预测时间上的变化，因为后者通常较小且可能受到在图像中更难以检测的变化驱动。来自马拉维和莫桑比克同一地点10年间隔的重复人口普查数据使我们能够评估模型是否确实可以从图像中提取出能够预测资产所有权时间变化的信息。

Results

国家一级财富预测的绩效

对于国家一级的财富预测，我们在每个模型上训练每个模型

国家单独进行，并对每个模型进行按国家五折交叉验证。CNN模型仅使用陆资卫星影像作为输入。XGBoost利用地理空间特征（地理特征）单独或与卫星影像统计特征结合使用。所有模型均训练以预测资产财富指数（AWI）^[29]。在马达加斯加、马拉维和莫桑比克，AWI 估算值在细粒度行政级别上生成并链接到影像数据。在布基纳法索，只有334个市镇的AWI 估算值可用（表[表号]）。¹ 这减少了布基纳法索训练数据的有效样本大小，如图右 panel 中聚类模式的调查测量资产财富所示。^{1 d}. This difference is reflected in the results. As shown in Figure 1 一个简单的变压器模型，在不一致地条件化地理特征的情况下，其他模型在马拉维、莫桑比克和马达加斯加的数据集中表现更佳。在这些国家，使用变压器模型进行的预测达到了² 当使用整个普查提取数据进行训练时，其值分别为0.83、0.70和0.62。在布基纳法索，由于有效样本大小较小，基于卫星影像和地理空间特征的XGBoost模型在这些模型中表现出最佳平均性能（解释了62.9%的变异）。仅使用卫星影像的简单变压器模型也保持竞争力（解释了57.4%的变异）。我们还限制训练样本数量为原始训练数据集的1%、5%、10%、25%和50%，以分析模型性能随训练样本大小的变化情况。根据这四个国家的结果，我们实证地识别出10%是一个关键转折点，在此之下，估计精度会迅速下降。

另一个减少财富预测训练样本采集成本的关键因素是每个样本中聚合的家庭数量。为了分析这一因素，我们随机从每个行政区域中抽取10户家庭来构建训练样本，从而为每个国家生成一个“10户家庭”的训练数据集。然后，我们在这个“10户家庭”的训练集中训练一个朴素的变压器模型，并仍然在原始完整的家庭测试集中评估其性能。结果（如图所示）^{1 b}) 表明我们的变压器模型使用10户家庭的数据进行训练时，其表现与使用所有家庭数据进行训练的模型相当。以马拉维为例，10户家庭的数据仅包括约23%的调查家庭。使用所有家庭数据训练的模型和使用10户家庭样本训练的模型之间的性能差距仅为3个百分点（82%对79%）。相比之下，当减少训练样本数量时，使用全部训练样本训练的模型与使用25%样本训练的模型之间的性能差距达到12个百分点（82%对70%）。

这些结果为高效的数据财富测量提供了重要的见解。当每张图像至少有10户家庭时，优先考虑对更多地理区域进行调查，而非在当前地理区域内调查更多的家庭，以使用我们的变压器模型预测财富。地理空间特征被认为是提高经济测量质量的重要辅助数据^[21]，在财富预测中广泛使用。在这里，我们为变压器模型设计了一个条件模块（详见“方法”部分），以实现地理空间特征和深度视觉特征的有效融合。结果如图所示。

1 C，建议地理空间特征显著改善

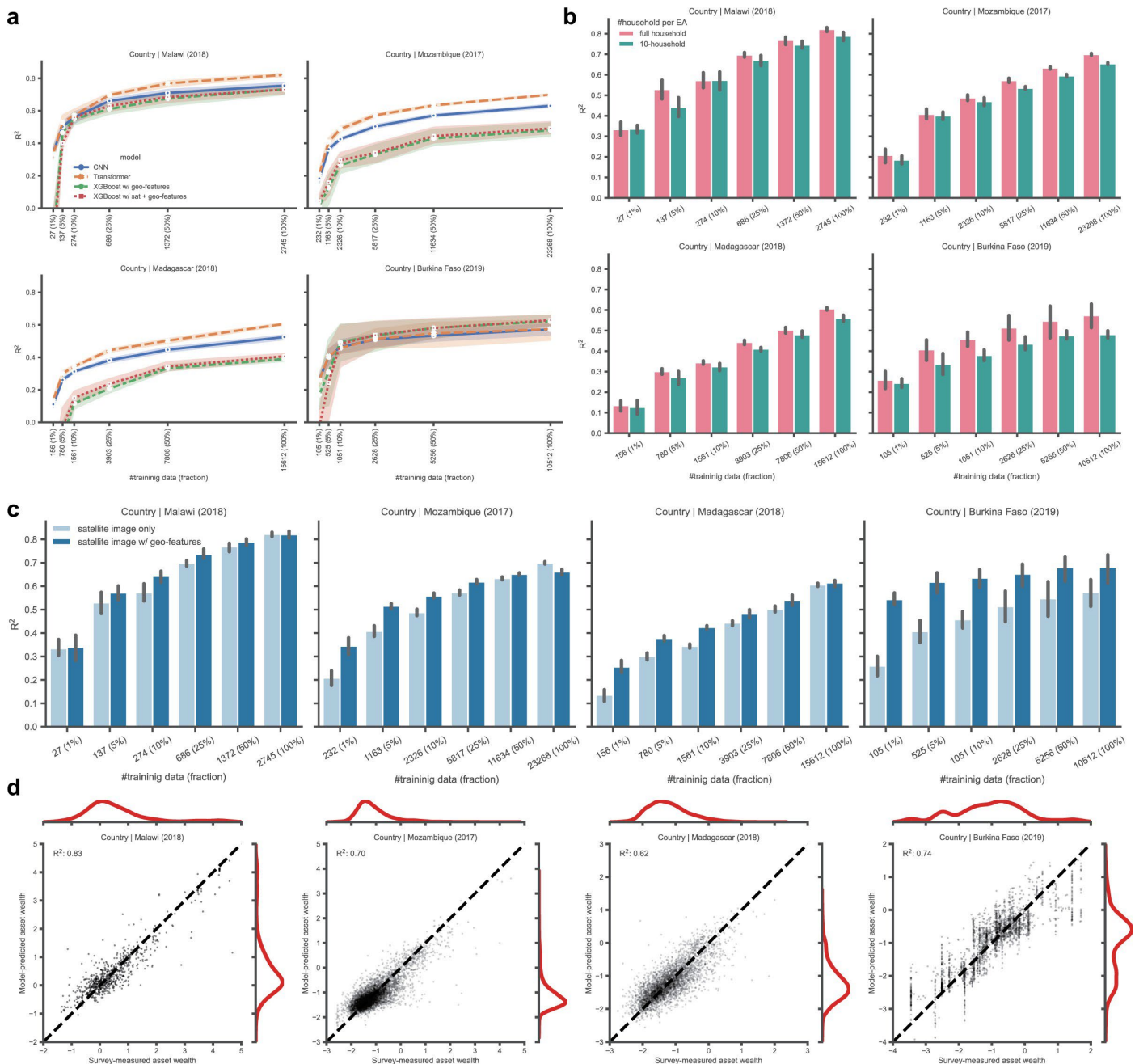


图 1：国家一级资产财富指数预测的表现。 性能比较：四个国家在不同机器学习模型训练下基于各种普查提取数据的不同表现。Negative R^2 值未显示。变分器结果不整合地理特征，并使用来自所有样本家庭的资产财富进行训练。**b.** 变分器模型在使用基于所有 Household 的资产财富与每个行政区域每十户家庭构建的资产财富进行训练后的性能对比。**c.** 具有和不具有集成地理空间特征的 Transformer 模型之间的性能比较。**d.** 调查测量的资产财富与表现最佳的预测财富的散点图。

模型在所有国家的表现均有所提升，特别是在布基纳法索，这主要是由于在 commune 级别将图像与调查数据链接时有效样本量较小所致。当训练样本量较小时，地理空间特征显得尤为有益，表明模型在小样本量下难以从原始图像中学习到最优的视觉表示，此时地理空间特征可以作为财富预测的重要补充。如图 1 中所示的方法，具有地理特征的变分器模型 (Transformer model with geo-features) 在图 1 中表现出最高的估计效果。² 在所有情况下，除了一个，当训练使用

莫桑比克的完整人口普查数据，在该数据中，朴素变分器模型略微更准确。图2中的四个网格财富地图，分辨率为4.8 km/像素，仅使用

我们的变分器模型和Landsat影像。无需进行地理空间特征准备，使用8块NVIDIA RTX A4000 GPU，整个制图过程可以在一小时内完成。这意味着我们的方法具有在国家范围内加速细粒度财富测量的巨大潜力。

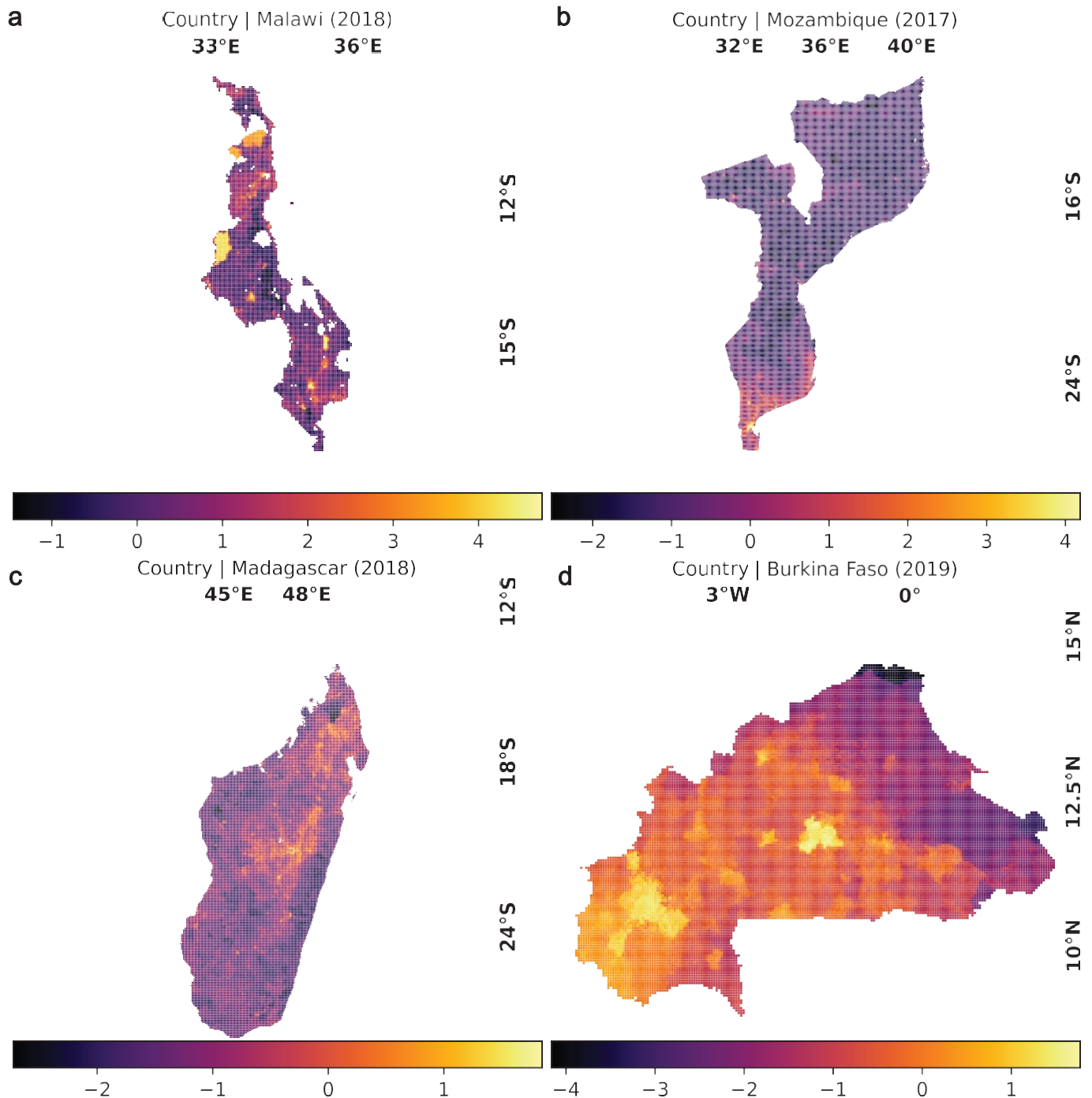


图 2：国家一级预测资产财富指数地图。 a. 2018 年马拉维国家级财富资产图。 b. 2017 年莫桑比克国家级财富资产地图 c. 马达加斯加 2018 年国家级财富资产地图。 d. 布基纳法索国家级财富资产地图（2019 年）。AWI 值来自国家特定模型，因此不具有国家间可比性。

预测国家一级财富变化的绩效

我们进一步通过五折交叉验证评估每个国家的财富变化预测。借鉴[29]的方法，CNN模型使用沿通道维度连接的双时相Landsat图像作为输入。我们的变压器模型单独处理每个双时相图像并通过共享权重的单图像编码器进行编码，并将编码特征进行拼接。

沿通道维度进行处理后再输入到最终的回归网络中。与之前的设置不同，XGBoost 只接受双时相卫星图像作为输入，因为没有可用的地理空间特征数据（莫桑比克在 2007 年和马拉维在 2008 年）。结果（图 3a）表明，基于完整样本训练的深度学习模型能够捕获马拉维 52% 的变化和莫桑比克 42% 的变化。当给定相同的输入数据时，深度模型的表现优于 XGBoost。这意味着代表权在

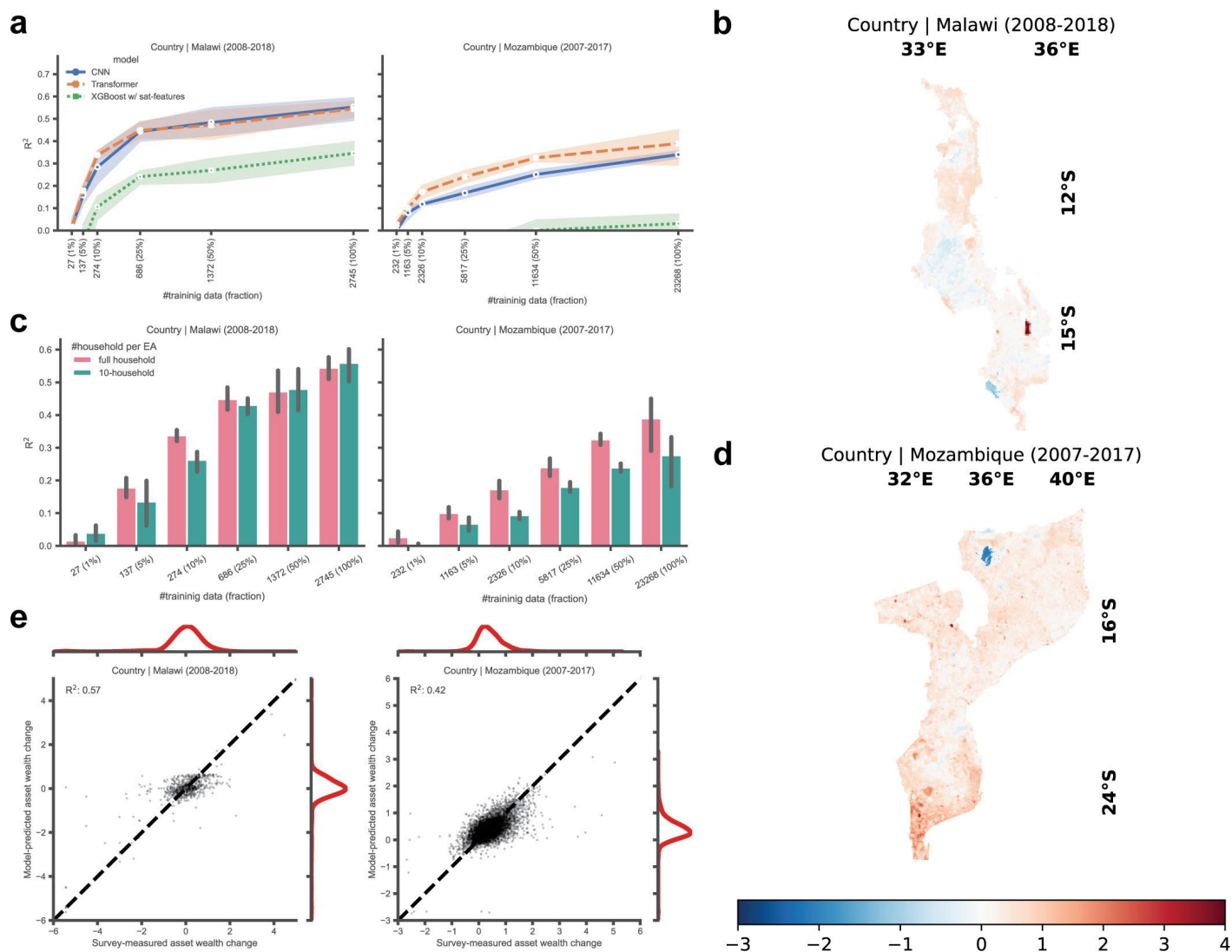


图 3：国家层面对资产财富指数年代际变化的预测表现。 两国在三种机器学习方法训练的不同人口普查提取数据比例下的性能比较。Negative R^2 值未显示。 **b.** 2008 年至 2018 年马拉维国家 / 地区财富资产变化图。 **c.** 基于每个行政区域所有 Household 和每个行政区域每 10 个 Household 构建资产财富的 Transformer 模型性能对比。 **d.** 莫桑比克 2007 年至 2017 年国家一级财富资产变化图。 **e.** 调查测量的资产财富与表现最佳的预测财富的散点图。

测量财富。我们的变压器模型在估计莫桑比克十年财富变化方面略微优于常用的 CNN，并且在马拉维方面达到了相当的性能。这种差异可能归因于训练样本大小和模型复杂度的变化。莫桑比克大约有 10 倍于马拉维的训练样本量，这可以解释为什么在莫桑比克，更具灵活性的变压器模型在估计方面优于 CNN。与上述横截面结果类似，我们模拟了两种数据稀缺性场景以预测变化：(i) 限制采样枚举区域的数量；以及 (ii) 将每个样本聚合的家庭数量减少到 10 个。结果（图 3 c）建议减少采样地点的数量比减少每批次样本的家庭数量对准确性的下降影响更大，这与实验结果一致。

国家层面的财富预测。

我们进一步证明了我们的变压器模型在马拉维年代际财富变化映射上的可扩展性（图 3 b）和莫桑比克（图 3 d）。与单一时间点的财富地图不同，双时间点的财富变化地图提供了更深入的经济动态见解，允许识别在过去一段时间内经历显著增长或衰退的地区。我们发现马拉维南部表现出更多的负面变化，表明过去十年间财富有所下降，而北部和一些中部地区则相对中性或略有积极表现。在莫桑比克，大多数地区整体上显示出财富增加的趋势，南部地区相比北部地区的财富增长更为显著。值得注意的是，在北部附近有一个蓝色区域，该区域的财富也有所减少。在两者中，

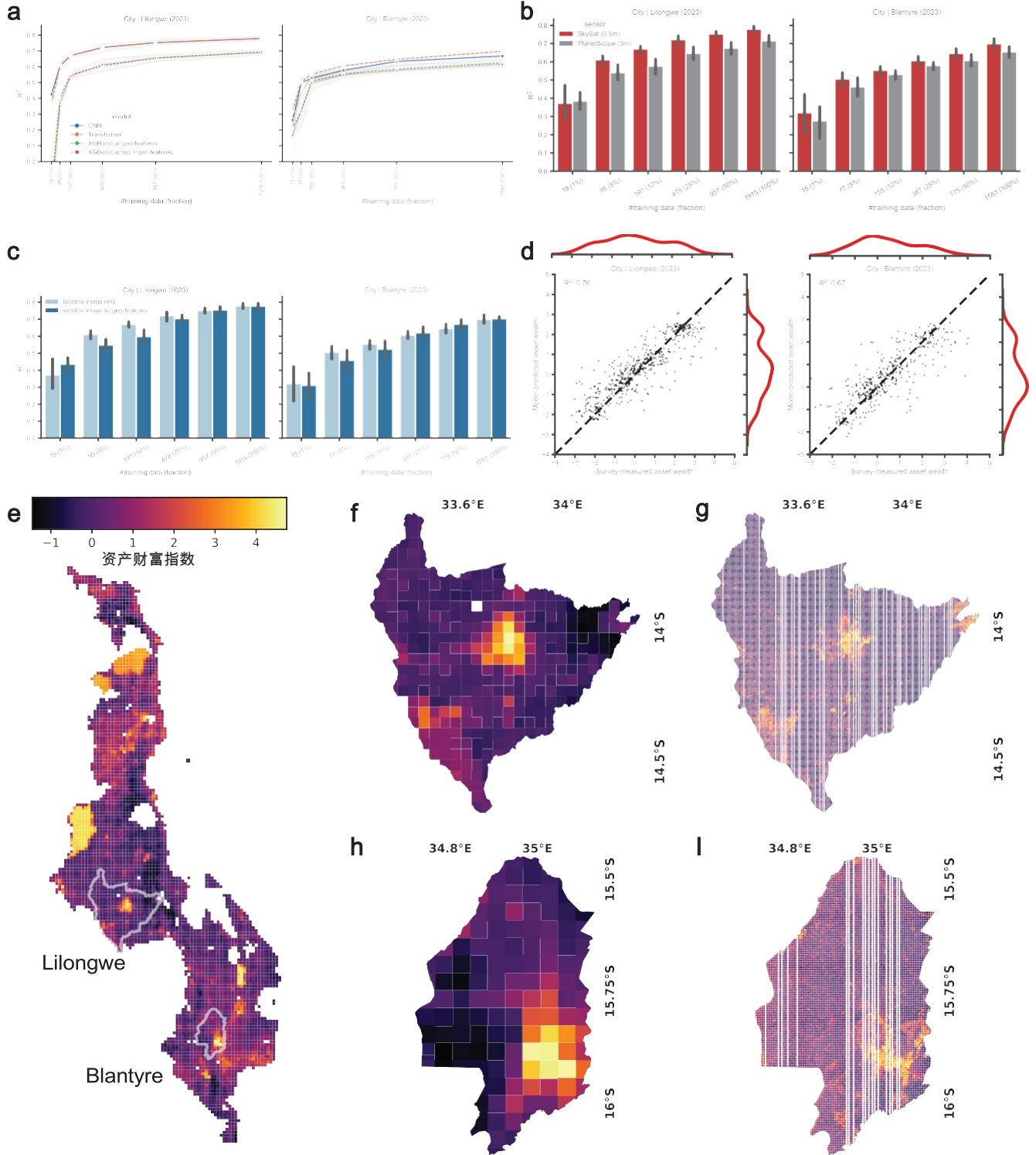


图 4：城市级资产财富指数预测的表现。 **a。** 两种城市在四种不同机器学习方法（训练数据分别为不同比例的人口普查数据）下的性能比较。 **b。** 使用 Skysat 图像和 Planetscope 图像估计的 Transformer 模型之间的性能比较 **c。** 具有和不具有集成地理空间特征的 Transformer 模型之间的性能比较。 **d。** 调查测量资产的散点图，预测资产来自表现最好的折叠。 **e。** 马拉维基于卫星的国家财富图。 **f。** 利隆圭国家级财富资产图。 **g。** 利隆圭市级财富资产图。 **h。** 布兰太尔的国家级财富资产图 **i。** 布兰太尔的市级财富资产图。

国家之间，财富分配的变化并不均匀，表明某些地区从经济增长中获益更多，而其他地区则落后。这种财富差距可能有助于了解经济政策、发展项目或外部因素（如气候变化影响）对这些变化的影响。

城市级财富预测绩效

基于Landsat的数据预测模型生成了空间分辨率为4.8 km的财富地图。对于某些应用，如城市内部的援助目标定位，更高分辨率的财富地图更为重要。为此，我们利用两个城市的家庭级人口普查数据，测试使用高分辨率卫星影像（PlanetScope和SkySat）进行财富预测。遵循上述两部分相同设置，我们评估了CNN、变压器和XGBoost模型，如图所示。4 a. 在两个城市中，CNN和Transformer模型始终优于XGBoost（无论是否包含地理空间特征）。这突显了高分辨率卫星图像提供的深层视觉表示对于城市级别财富预测的重要性，这是基于深度学习的模型与XGBoost之间的主要差距。我们的Transformer模型在Blantyre的表现明显优于CNN，尤其是在使用完整数据集时；同时，在Lilongwe中，它也实现了与CNN相当的性能。在两个城市中，所有模型随着训练数据比例的增加而表现出显著的改进；然而，当训练数据比例达到约25%-50%后，性能提升开始停滞，因此在此阈值以上时，改进幅度逐渐减小。

我们比较了两种常用的高分辨率多波段卫星图像，即PlanetScope（3米分辨率）和SkySat（0.5米分辨率），如图所示。4 b. 结果表明，SkySat在各种训练数据比例下均一致地优于PlanetScope，这一现象在两座城市中均有体现。两种传感器均捕获4波段（红、绿、蓝和近红外）卫星影像，但它们在空间分辨率和扫描宽度上存在差异。这突显了在使用我们的变压器模型准确测量财富时，城市空间细节的重要性。尽管PlanetScope的平均性能低于SkySat，其更宽的扫描宽度和更高的重访频率使其能够提供更为全面的卫星影像，从而有助于大规模财富测绘并保持较高的准确性（如图所示）。4 SkySat适用于高精度要求的当地区域财富测量，而PlanetScope更适合进行大规模财富mapping，以获得宏观洞察。

尽管城市级别的财富预测前景可观，但训练城市级别的变压器仍然需要大量的样本，而这些样本的收集成本较高。我们还评估了整合地理空间特征是否能够减少所需的训练样本数量。

4 c，我们发现城市级别的财富预测。如图所示，与国家级别预测不同，地理空间特征通常会降低模型性能。这是因为地理空间特征总是从低分辨率的卫星影像中提取的，例如Landsat（30米分辨率），以实现全球覆盖。这些低分辨率的地理空间特征会引入到基于高分辨率卫星图像的财富预测中的空间误差。例如，一个粗略的土地覆盖图可能会忽略一些虽小但重要的区域。

城市地区的空间细节。在这种情况下，我们的变压器模型可以从遥感图像中学习更准确的财富表示，即使训练数据样本大小不同也是如此。

总体而言，我们在马拉维的两个城市——利隆圭和布兰太尔——实现了精确的大规模、城市级财富地图绘制。与分辨率高达4.8公里的国家层面财富地图（如图所示）相比，我们展示了准确的城市级财富映射。4 f和4 h），我们的0.3公里分辨率城市级财富图（图4 g和4i）提供了这两个城市前所未有的精细的空间财富分布，并且表现出强大的性能，能够解释利隆圭大约76%的变化情况，解释布兰太尔大约67%的变化情况（如图所示）。4 d）。

Discussion

这篇论文提出并评估了使用视觉变换器架构来解决结合调查数据和卫星数据以在细粒度空间尺度上生成财富估计值的多个开放性问题。对于横截面财富预测，通过对四个国家地理参考的人口普查摘录进行仔细评估表明，变换器模型的估计效果良好。当与陆卫星像（Landsat imagery）结合使用时，R² 变压器模型在包含地理特征的情况下，对于资产指数预测的表现优于常用于此类任务的CNN和XGBoost模型，在四个国家的所有样本大小中均表现优异。在所有国家中，当训练数据少于10%的普查提取时，准确性会急剧下降。在莫桑比克和马达加斯加，使用变压器模型生成的估计值能够解释财富变化约20至30个百分点以上的变异，而使用XGBoost和地理特征生成的估计值则无法达到这一水平。在所有四个国家中，变压器模型也优于CNN，在马达加斯加和莫桑比克，这种优势甚至可达约5个百分点。在莫桑比克、马达加斯加和布基纳法索的小样本情况下，将地理特征纳入变压器架构可提高性能5至10个百分点。在预测变量时，变压器模型同样表现出色。

以0.3公里的尺度在城市内移动，实现²在利隆圭和布兰太尔分别达到了0.76和0.67。然而，在这个尺度下纳入地理特征降低了模型性能，因为这些特征是从较低分辨率的Landsat图像构建的。最后，transformer模型在预测莫桑比克和马拉维的财富指数十年变化²比CNNs和XGBoost生成了更准确的估计值。模型预测实现了²在精细的空间级别下，马拉维的价值为0.57，莫桑比克的价值为0.42，尽管缺乏可用的地理特征。这相较于之前的0.15至0.17有了显著的提升。报告人²⁹在可比的环境中进行，并展示了在有足够的训练数据的情况下，将变压器模型与图像结合以在细粒度级别估计财富变化的可行性。结果表明应用²ypsum变换器的优势。

利用地理空间特征生成贫困和财富高分辨率预测的模型。这反过来强调了开发工具、文档和培训材料的重要性，以使估计工作能够在国家统计局、国际组织和其他数据提供商中实现。此外，开发和评估与预测相关的不确定性估计方法对于促进实施至关重要。

结果还强调了获得

大量训练数据以估计预测模型。通常而言，当用于训练模型的图像数量低于总体的10%时，预测性能会迅速恶化。然而，结果对限制用于生成训练数据标签的样本大小更加稳健。此外，在布基纳法索，预测精度保持较高水平。

尽管由于将卫星图像与更高地理级别的调查数据链接起来，训练数据的有效样本大小有所减少。未来的研究可以探讨在使用类似家庭调查通常收集的小样本量训练变压器模型时进一步提高性能的方法。

最终，研究结果表明，使用变压器模型预测财富和家庭福祉变化具有潜在价值。未来的研究可以探讨变化模型中的参数在时间上和/或空间上的稳定性。这可能为利用地理空间数据生成福利变化的近似微观估计指明方向，在缺乏调查数据的情况下尤其如此。

Methods

我们描述了大型多分辨率和多时相贫困数据集的详细情况，以及贫困及其变化预测方法和评估方法。

多分辨率和多时态贫困数据集。

我们利用四个低收入国家（马拉维、莫桑比克、布基纳法索和马达加斯加）以及马拉维境内两个城市的资料作为研究区域，位于非洲。这些国家被选中是因为可以获得地理位置标识符，从而可以在现有的普查摘要中进行选择。这使我们能够将模型应用到迄今为止最全面的规模上，调整通用模型以适应特定国家甚至城市，并且能够在三个时空场景下稳健地模拟数据稀缺对模型性能的影响：国家层面的财富水平预测、十年国家层面的财富变化预测以及城市层面的财富水平预测。

资产财富指数 (AWI)。 我们使用国家普查问卷的数据构建资产财富指数。我们在莫桑比克和马达加斯加利用完整的普查数据，在马拉维和布基纳法索利用普查摘录，从而形成了一个包含四个国家超过1200万户家庭的总数据集，每个国家的数据集中有超过70万户家庭。相比之下，以往的研究中，²⁹，²³ 利用DHS数据涵盖了大约23个非洲国家的500,000户家庭，而LSMS数据则调查了约9,000户家庭，提供了全面的描述。

1 p
数据集的表大小。

从人口普查问卷中，我们将七种住房特征（房屋类型、墙体材料、屋顶材料、地板材料、水源、厕所类型和能源来源）按1到5的等级进行排序。此外，我们使用二元分类（拥有/未拥有）来评估六项资产（收音机、电视机、固定电话、汽车、摩托车和自行车）的存在情况。然后对这些数据进行标准化处理，并用于构建主成分分析（PCA）模型，从中提取出第一个主成分作为资产财富指数。¹³，²⁵，²⁹。然后，将资产财富指数标签从 household 汇总到

表 1
全国人口普查数据详细信息

Country	# Images	# Adminareas	# Households Surveyed
马拉维 (2008)	3,432	12,412	572,764
马拉维 (2018)	3,432	18,700	796,925
莫桑比克 (2007 年)	37, 325	45,244	4,797,372
莫桑比克 (2017) 37, 325		67,218	6,119,847
布基纳法索 (2019 年) 13, 142		334	875,872
马达加斯加 (2018) 31, 182		14,328	4,518,322

在行政区域级别进行标注，然后根据每个像素所属的行政区域对其进行标注。最后，通过平均像素级别的资产财富指数地图来获得每个图像的真实值（ground truth）。由于为每个国家构建了不同的主成分分析（PCA），因此一个国家中的0值并不对应另一个国家中0值相同的财富水平。

卫星图像。 对于国家层面的贫困及其变化预测，我们收集了马拉维、莫桑比克、马达加斯加和布基纳法索的白天 Landsat 卫星影像（每像素 30m）。马拉维和莫桑比克拥有双时间点影像对。我们的 Landsat 影像数据集是基于每个国家普查年的 3 年中值无云像素构建的。对于有两个普查时期的国家，早期普查使用 Landsat 5 和 Landsat 7 的影像，而近期普查则使用 Landsat 8 的影像。每张 Landsat 图像固定大小为 150×150 像素，每幅图像覆盖面积为 20.25 平方公里。² (4.5×4.5km² 这些图像包含六个波段，分别为红色、绿色、蓝色、近红外、短波红外1和短波红外2。为了预测城市级别的贫困情况，我们收集了PlanetScope（每像素3米）和SkySat（每像素0.5米）多光谱卫星图像，以覆盖利隆沃和布兰太尔地区的每个行政区域。基于行政区域的平均大小，我们实证定义每个网格的大小为0.3×0.3公里。²，每张PlanetScope图像包含100×100像素，每张SkySat图像包含600×600像素。尽管使用了2018年的户级人口普查数据，但我们利用了2023年4月获取的可用的PlanetScope和SkySat影像。这些图像包含红、绿、蓝以及近红外波段。

地理空间特征。 此外，我们补充使用公开可用的处理地理空间特征的卫星影像，我们称之为 *地理特征*。这些地理特征涵盖了人口、发展和环境统计数据。这些特征包括人口结构^[18]、人口密度^[28]，年降雨量^[1]，最低和最高温度^[1]，夜间灯光^[3]，^{11]}，terra 净初级产品^[10]，aqua 净初级生产^[24]，手机塔计数^[22]、不透水表面变化年 L^[26]，建筑计数^[27]，建筑物^{5]},GHS^[15]，土地覆盖类型^{[地区[27]}、土壤酸碱度^[16]和土壤有机碳^[16]。图中提供了一个数据点的可视化表示⁵；请注意，最终资产财富标签被组合成一个标量值。

培训贫困衡量模型

比较包括基于树的模型，即极端梯度提升 (XGBoost)^[7]和两种先进的深度学习模型（基于卷积神经网络的模型和基于编码器-线性架构的变换器模型^[29]）。根据经验数据

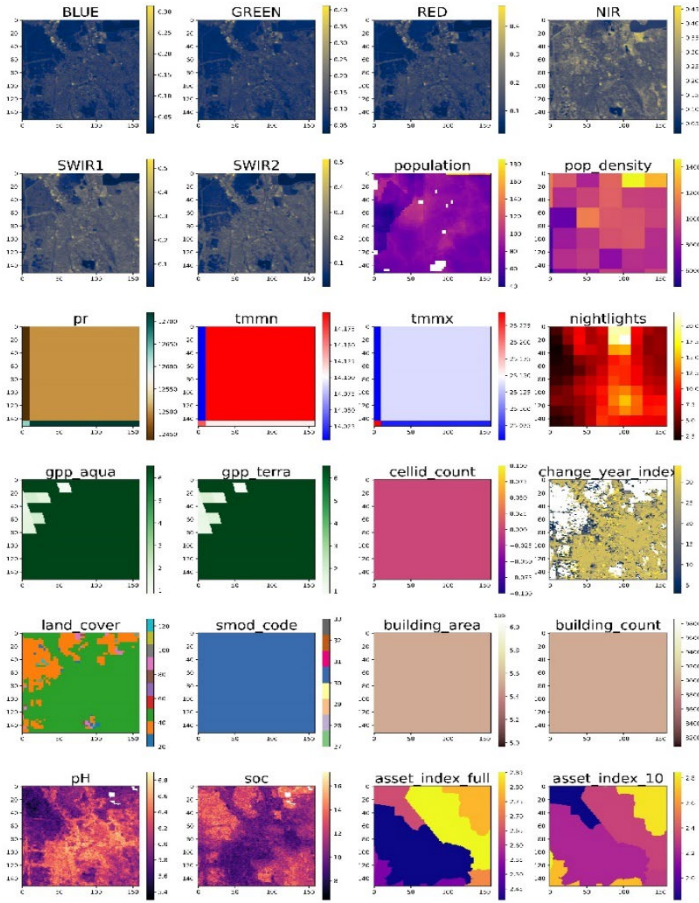


图 5: 一个带有地理空间特征和资产财富指数标签的卫星图像示例。这是一名级训练样本，针对的是整个国家。

从 [14] 的系统观测，我们选择 SwinV2 - T [19] 作为变压器模型的代表性骨干。

XGBoost。 一种XGBoost回归模型被用于本论文中。通过实验，我们确定将图像级通道矩直接作为XGBoost输入特征时性能最佳。仅使用卫星影像时，我们使用了三个矩（均值、标准差和偏度）；而结合使用卫星影像和地理空间特征时，我们仅使用了一个矩（均值）。对于财富变化预测，我们将两个年份的所有通道矩输入同一个XGBoost模型，该模型经过训练可以直接预测AWI的变化。

卷积神经网络 (CNN)。 以下 [29],我们构建了一个基于ResNet-18的CNN模型进行财富水平预测和变化预测。对于财富水平预测，我们首先使用ResNet-18提取深层特征，然后通过全局平均池化层计算嵌入向量。在最后一个深层特征之后附加了两个多层感知器（MLP）层以预测AWI。对于贫困变化预测，主要的区别在于特征提取。我们将两个时间点的图像沿通道轴连接起来，并将结果输入到CNN中以提取深层特征。

变压器架构，我们首先采用SwinV2-T作为骨干网络以提取深层层次特征。与CNN类似，我们附加了两层MLP层来预测AWI。为了将地理空间特征整合到该变压器模型中，我们提供了一种条件机制，采用了标准的交叉注意力层，以在可学习的方式中将地理空间特征融入到深层特征中。SwinV2-T生成四个深层层次特征；因此，我们采用了四个交叉注意力层来进行条件处理。通过四次条件处理，最终的深层特征能够很好地与地理空间特征整合。最后的深层特征基于上述两层MLP层用于财富回归。对于贫困变化预测，我们采用了一种Siamese网络架构，在双时间图像中共享一个SwinV2-T骨干网络，即我们独立地为每张图像使用SwinV2-T提取深层特征。然后我们将这两个深层特征集合沿通道轴进行连接，并将结果张量输入到两层MLP层中以预测贫困变化。

深层模型的实现细节。 我们使用相同的配置训练所有深度模型。所有模型均通过最小化均方误差损失并使用AdamW优化器[20]从头到尾进行端到端训练。每个模型训练100个周期。总批次大小为32，学习率为常数1e-4，权重衰减为1e-2。训练数据增强采用D4旋面群变换以缓解过拟合。使用全连接（MLP）层来预测AWI。对于贫困变化预测，主要区别在于特征提取。我们将两个时间点的图像沿通道轴拼接，并将结果输入卷积神经网络（CNN）以提取深层特征。

模型评估

数据分割和交叉验证。 为了确保对模型性能进行稳健的评估，我们采用了五折交叉验证的方法，为每个国家或城市训练了五个不同的模型。每个模型在四折的数据上进行训练，并在剩余的一折上进行测试。通过均匀采样行政区域来创建折分割，确保同一行政区域内的所有图像都在同一个折中，从而使五个折中的图像数量大致相同。 R^2 用作级别和更改预测的度量。

模拟数据稀缺性。 为了调查模型在数据稀缺情况下的性能，我们模拟了两种数据可用性受限的场景。**(i)** 限制图像数量。**我们减少了用于训练数据的采样图像的数量。在限制数据时，我们在每个四个训练折叠中均匀和随机地采样图像。我们使用了完整训练集样本的1%、5%、10%、25%、50%和100%进行了实验。**(ii)** 限制图像中的户数。**我们引入了一个替代的“10户资产财富”标签。对于10户资产财富（AWI）标签的创建，从每个区划区域中均匀地抽样家庭。之后，10户资产财富标签的创建与上述全量资产财富标签的创建方式相同。这些标签仅使用每个区划区域内的10户家庭的数据，而我们的全量资产财富标签则使用成百上千户家庭的数据。

Acknowledgments

我们感谢Lina Cardona、Carlos Da Maia、Francis Mulangu、Mario Negre、Soudiki Soubeiga和Michael Weber在获取数据方面提供的帮助；感谢Haishan Fu、Olivier Dupriez、Craig Hammer和Jed Friedman提供的支持；以及Brian Amaro、Nahum Maru和Rohan Sikand在初步分析方面的帮助。本项目部分资金由Knowledge for Change计划第四阶段资助的研究项目“理解低收入和中等收入国家区域内经济福祉差异的趋势”以及Keck基金会提供。

作者贡献

DN、TK、MB、SE和DL构思了该项目并设计了分析方案；RL处理了数据；ZZ和TW负责了分析工作；ZZ、DL、DN和MB撰写了论文。

代码和数据可用性

进行分析和生成数字的代码可在。我们不

https://github.com/Z-Zheng/dynamic_highres_pov_erty

目前获得了国家统计局的许可，可以分享 Household Level 数据或 Image-Level 标签。

参考文献

[1] 亚巴佐格卢, J.T., 托布罗夫斯基, S.Z., 帕克, S.A., 赫格维希, K.C., 2018. Terraclimate , 1958-2015年全球高分辨率月度气候及气候水分平衡数据集. 科学数据 5, 170191. doi: 10.1038/sdata.2017.191 . [2] 阿约什, K., 穆肯特, B., 布克, M., 洛尔, D., 埃蒙, S., 2020. 使用卫星图像中的目标检测生成可解释的贫困地图. arXiv预印本 arXiv:2002.01612 . [3] 巴格, K., 埃尔维德吉, C.D., 吉什, T., 西金, D., 2010. 利用DMSO-OLS数据开发2009年稳定灯光产品, 在《亚太先进网络会议论文集》第30卷中, 第114页。 [4] 布伦斯特龙, J., 卡达穆罗, G., 俄恩, R., 2015. 从移动电话元数据预测贫困和财富. 科学 350, 1073-1076. [5] 舒尔霍恩, M., 利西, M., 特森巴扎尔, N.E., 赫尔德, M., 贝尔茨, L., 斯梅茨, 克普勒斯全球土地覆盖图层—集合2. 远景 B , , 2020 年. Sensing 12. URL : <https://www.mdpi.com/2072-4292/12/6/1044> , doi: 10.3390/rs12061044 . [6] 布尔ke, M., 德里斯科尔, A., 洛尔, D.B., 埃蒙, S., 2021. 利用卫星影像理解并促进可持续发展. 科学 371, eabe8628. [7] 陈, T., 格雷森, C., 2016. XGBoost : 一种可扩展的树增强系统 , 在第22届ACM SIGKDD国际知识发现与数据挖掘会议上发表 , 页码 : 785–794. [8] 金, G., 方, H., 查特吉, S., 布鲁门斯托克, J.E., 2022. 全球低收入和中等收入国家的微观财富估计. 美国国家科学院院刊 119, e2113658119. [9] 丹, H.A.H., 赛拉朱丁, U., 2020. 追踪可持续发展目标 : 新兴测量挑战及进一步反思. 发展 127, 104570. [10] 迪丹, K., 2021. MODIS/Terra植被指数16天L3全球500米sin网格v061 [数据集]. URL: <https://doi.org/10.5067/MODIS/MOD13A1.061> . 2024年6月8日访问。 [11] Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS夜间灯光。《遥感国际杂志》38, 5860–5879. [12] Engstrom, R., Hersh, J., Newhouse, D., 2022. 贫困从太空视角看 : 利用高分辨率卫星图像估算经济福祉。《世界银行经济学评论》36, 382–412. [13] Filmer, D., Pritchett, L. H., 2001. 无支出数据或眼泪下的财富效应估计 : 印度各邦教育入学率的应用。《人口学》38, 115–132. [14] Goldblum, M., Sourì, H., Ni, R., Shu, M., Prabhu, V.U., Somepalli, G., Chattopadhyay, P., Ibrahim, M., Bar des, A., Hoffman, J., Chellappa, R., Wilson, A.G., Goldstein, T., 2023. 骨干网络之战 : 跨计算机视觉任务大规模比较预训练模型, 在“第37届神经信息处理系统数据集和基准跟踪”。 [15] Gong, P., Li, X., Wang, J., Bai, Y., Chen, B., Hu, T., Liu, X., Xu, B., Yang, J., Zhang, W., Zhou, Y., 2020 . 1985年至2018年间全球人工不透水面积年度地图 (GAIA) 。《遥感环境》236, 111510. URL: <https://www.sciencedirect.com/science/article/pii/S0034425719305292> ,doi: <https://doi.org/10.1016/j.rse.2019.111510> . [16] 亨格, T., 米勒, M.A.E., 克里赞, J., 莱谢特, K.D., 斯里拉, A., 卡利巴达, M., 安托伊耶维奇, O., 格卢西卡, L., 道GAN, I., 舒特查, M.N., 利纳斯, J.G.B., 狼, J.W., 梵登布什, R., 基尔彭, B., 德朱斯, J.M., 里贝罗, E., 马克米伦, R.A., 2021. 使用两尺度集成机器学习方法绘制的30米空间分辨率的非洲土壤性质和养分分布. 科学报告 11, 6130. doi: 10.1038/s41598-021-85639-y . [17] 李安 (Jean) , 伯克 (Burke) , 谢 (Xie) , 戴维斯 (Davis) , 洛贝尔 (Lobell) , 埃蒙 (Ermon) , 2016. 结合卫星影像与机器学习预测贫困. 科学 353, 790-794. [18] 林德 (Linard) , 吉尔 (Gilbert) , 诺伍德 (Snow) , 塔特姆 (Tatem) , 2012. 2010年非洲的人口分布、聚落模式和可达性. 公共科学图书馆 综合 7, e31743. doi: 10.1371/journal.pone.0031743 . [19] 刘泽, 胡海, 林洋, 袁昭, 谢壮, 魏宇, 宁嘉诚, 曹一, 张子昂, 董乐, 等., 2022. Swin Transformer V2: 扩容能力和分辨率, 在: IEEE/CVF计算机视觉与模式识别会议论文集, 页码: 12009–12019. [20] 洛什奇洛夫, I., 2017. 分解权重衰减正则化. arXiv预印本 arXiv:1711.05101. [21] Newhouse, D., 2024. 使用地理空间数据估算贫困和财富的小区域估计 : 迄今为止我们学到了什么? 《加尔各答统计协会通报》76, 7–32. [22] OpenCellID, 2024. Opencellid : 全球最大的开放基站数据库。访问日期 : 2024-06-08. URL : <https://opencellid.org> . 数据遵循 Attribution-ShareAlike 4.0 International (CC BY-SA 4.0). [23] Pettersson, M.B., Kakooei, M., Ortheden, J., Johansson, F.D., Daoud, A., 2023. 卫星图像的时间序列提高非洲街区层面贫困程度的深度学习估计值, 在IJCAI上 , 第6165–6173页。 [24] Running, S., Zhao, M., 2021. MODIS/Aqua净初级生产力缺口填充年度L4全球500m SIN格网v061 [数据集]. URL: <https://doi.org/10.5067/MODIS/MYD17A3HGF.061> . 查阅日期 : 2024-06-08. [25] Sahn, D.E., Stifel, D., 2003. 探索在缺乏支出数据时的替代福利衡量方法。《收入与财富评论》49, 463–489. [26] Schiavina, M., Melchiorri, M., Pesaresi, M., 2023. GHS-SMOD R2023A - GHS 结算图层, 基于城市化程度方法 (阶段 I) 对 GHS-Pop R2023A 和 GHS-Built-S R2023A 的应用, 多时段 (1975-2030) 。 URL: <http://data.europa.eu/89h/a0df7a6f-49de-46ea-9bde-563437a6e2ba> , doi: 10.2905/AODF7A6F-49DE-46EA-9BDE-563437A6E2BA . [数据集] [27] Sirko W., Kashubin S., Ritter M., Annkah A., Bouchareb Y.S.E., Dauphin Y.N., Keyser D., Neumann M., Cissé M., Quinn J., 2021. 大规模范围内的建筑物检测 : 基于高分辨率卫星影像. CoRR abs/2107.12283. URL: <https://arxiv.org/abs/2107.12283> , arXiv: 2107.12283 . [28] WorldPop, 地理与环境科学学院, 南安普敦大学 ; 地理与地球科学系, 路易斯维尔大学 ; 地理学系, 纳慕尔大学 ; 国际地球科学信息网络中心 (CIESIN) , 哥伦比亚大学, 2018. 全球高分辨率人口分母项目 URL: - 由比尔和梅林达盖茨基金会(opp1134076) 资助。 <https://dx.doi.org/10.5258/SOTON/WP00674> . 查阅日期 : 2024-06-08. [29] YeH, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D., Ermon, S., Burke, M., 2020. 使用公开可用的卫星图像和深度学习理解非洲的经济福祉. 自然通讯 11, 2