

SCIENTIFIC REPORTS

开放

使用深度学习和 street 图像测量社会、环境与健康不平等

将有关 1,2,3, John W. polak^{3,4}, James E. bennett^{1,2} & Majid Ezzati^{1,2,5}

收稿日期:2018年11月14日
录用日期:2019年3月18日

发布日期:18 April 2019

城市是世界上越来越多的人的家园。目前,这一数据很难追踪高时空性城市的社会、经济、环境与健康结果。我们应用了深度学习的方法用于测量收入、教育、失业率、住房、生活环境、健康和犯罪。我们的模型直接从预测了不同的结果无需提取中间用户定义的特征的原始图像。来评估的性能方法,我们首先使用地面真实数据在伦敦的图像子集上训练神经网络来自官方统计数据的高空间分辨率。然后,我们比较了训练过的网络是如何分离的在未用于训练的图像中,最好的和最差的十分位数用于不同的结果。最好的居住环境和平均收入两项表现最佳。分配最少成功的犯罪和自我报告的健康(但不是客观衡量的健康)。我们还评估了在伦敦培训的网络如何预测 UK 的其他三个主要城市:伯明翰,曼彻斯特和利兹。可转让性分析显示,在伦敦培训的网络, fine-只对其他城市 1% 的图像进行了调整,获得了与训练过的图像相似的表现来自目标城市本身的数据。我们的研究表明,街道图像具有潜力补充传统的基于调查和行政数据源的高分辨率城市监测以衡量不平等现象,并监测旨在解决不平等问题的政策的影响。

估计有 42 亿人居住在城市,占世界人口的 55%¹。在欧洲,5.52 亿城市人口占总人口的 75%¹。与农村居民相比,城市人口通常拥有更高的平均经济地位和教育水平,平均健康状况也更好²⁻⁴。与此同时,城市内部在收入、教育、生活和社区环境、健康和安全性方面存在着严重的不平等^{2,3,5-8}。在伦敦这样的大城市,不平等现象尤其严重,穷人和富人生活在一起,他们的生活环境和健康状况存在巨大差异⁸⁻¹¹。减少不平等是全球可持续发展议程的首要任务¹²在伦敦等城市,这是一个政策目标^{13,14}。

从收入、健康、教育、就业、犯罪、住房和生活环境等福祉的不同维度考虑,城市不平等具有复杂的空间分布,收入和就业等一些指标在空间上重叠,而犯罪和过度拥挤等其他指标则具有不同的空间格局。城市景观和不平等也会因士绅化、城市衰落和城市经济活动变化等过程而发生变化。以高空间和时间分辨率衡量城市不平等现象对于提供信息和评估政策至关重要。少数国家拥有完全关联的数据集¹⁵⁻¹⁷这些数据能够实时衡量卫生、社会和经济状况方面的不平等^{18,19}。然而,如此丰富的数据是罕见的,在大多数国家和城市,衡量城市不平等是一项重大挑战,因为关于不同社会、环境和健康措施的数据来自不同的来源,具有不同的空间分辨率和频率,并且往往使用昂贵的过程收集。

通过公共、私人 and 众包数据越来越多地获得大规模图像数据,这些数据有可能显著提高我们测量城市的速度、频率和本地化程度

¹ 伦敦帝国理工学院公共卫生学院, UK 伦敦。² MRC-PHE 环境与健康中心, 伦敦帝国学院, UK 伦敦。³ 伦敦帝国理工学院城市系统实验室, 伦敦, 伦敦, SW7 2AZ。⁴ 伦敦帝国理工学院土木与环境工程系交通研究中心, 伦敦, SW7 2AZ。⁵ 世界卫生组织非传染性疾病监测和流行病学合作中心, 伦敦, UK。通信和材料请求应发送至 E.S.(电子邮件: esra.suel@imperial.ac.uk)

特性²⁰⁻³⁰。图像可以用来衡量城市地位及其居民的福祉,这是以一系列现象为前提的:首先,城市和城市生活的一些特征,如住房质量和居住环境,以建筑材料和年久失修、空气和噪音污染源以及绿地的形式有直接的视觉信号³¹⁻³⁵。其他因素,如贫困,可能是可见的,因为它们影响或与住房和社区环境、人们使用的车辆类型,甚至商店类型等特征相关^{32,34-38}。最后,由于健康和犯罪等结果与社会和环境因素的关系,即使它们的绝对水平可能更难预测,但在图像中也可以检测到它们在城市内部的差异^{28,29,39-41}。这种潜力反映了历史上使用视觉检查和信号系统推断社会经济和环境特征的尝试。例如,查尔斯·布斯(Charles Booth)在19世纪90年代绘制伦敦的贫困地图时,依赖于对社区的面对面目视检查,并使用了街道、住宅和居民服装等指标³³。

使用人工视觉检查技术和机器学习方法,越来越多地将街道图像作为经济学和环境健康方面的信息来源进行调查^{24-29,36,42}。其他研究试图衡量感知的邻里属性,如安全性和社会阶层²⁴⁻²⁷,房价,犯罪率,以及街道图像中的人口密度^{42,43}。这些研究中的每一项都使用了来自不同环境和城市的一组不同的图像,采用了不同的分析方法,并在不同的空间尺度上预测了一组不同的结果。然而,在同一城市或国家,使用相同的方法,相同的图像,在相同的空间分辨率下,还没有对可能或可能在图像中检测到的多种结果进行可比评估。由于缺乏可比性分析,很难评估我们是否可以以及在多大程度上使用图像来全面地检测城市不平等。

在这里,我们以伦敦为例,评估了使用深度学习的公开街道图像来衡量人类福祉多个方面的不平等的可行性和性能,包括收入、健康、教育、就业、犯罪、住房和生活环境。除了其新颖的经验范围-多个结果中不平等的可比分析之外,我们的方法还进行了两项方法创新:首先,我们直接从原始图像中预测感兴趣的结果,而不提取中间用户定义的特征⁴⁴。网络自动使用与测量任务相关的特征,而无需研究人员明确指定相关的预定义特征(例如屋顶类型、树木、车辆)。因此,我们的方法消除了提取预定义特征的中间步骤,这一过程需要在可能无法公开的数据库上进行单独的训练^{42,43}。在视觉识别任务中,自动学习相关特征的深度学习方法通常倾向于优于基于手工特征的方法⁴⁴并且不依赖于研究人员预定义的特征是否与手头的任务相关,以及在多大程度上相关。其次,我们研究了根据伦敦数据训练的模型是否可以转移到其他城市,这反过来表明了与福祉指标相关的视觉特征在多大程度上在各个城市之间共享。对可转移性的评估还显示,在拥有高分辨率数据的城市上训练的网络是否可以用于其他数据较差的城市的监测。

社会、环境和健康结果。我们使用了12个结果,用于训练和测试网络的数据可从政府统计数据中获得,用于低层超级输出区域(LSOA;平均人口数为1614)。对于每个城市的每一项结局,我们计算了LSOAs的十分位数,十分位数1对应最糟糕的10%,十分位数10对应最好的10%(图1)。结局标签及其详细定义列于下文。有关结果的数据来自英国人口普查⁴⁵以及英国贫困指数⁷对于那些衡量剥夺的指数。剥夺指数往往集中在福祉结果分布的低端,因此没有捕捉LSOAs与富裕之间的差异。我们如何计算每个结果指标的细节如下所示。

- 生活环境剥夺:剥夺指数计算使用当地空气质量估计,每1000名居民的交通事故率,住房条件差,定义为不符合UK体面的住房标准,以及缺乏中央供暖的住房,作为衡量住房供暖昂贵的措施。基础数据来自国家统计局(包括住房调查)、UK航空信息资源和交通部。
- 平均收入:大伦敦当局家庭平均收入估计(仅适用于伦敦)。
- 收入剥夺(Income deprivation):与失业或低收入导致的低收入有关的剥夺指数。基础数据包括政府援助计划的索赔人数据集。
- 入住率:被归类为过度拥挤的家庭的百分比。至少比需要的房间少一间)。ONS根据家庭成员的年龄及其彼此之间的关系,使用标准公式得出所需的房间数量。
- 住房和服务障碍:基于住房和当地服务的物质和经济可及性的剥夺指数。ONS用于计算这一指标的数据包括无家可归、负担能力、过度拥挤以及到学校、超市、医疗服务和邮局的距离。
- 教育:没有接受过二级教育的人的百分比。最高获得资格的五个类别是:无资格、1级、2级、3级和4级及以上。英国的资格等级由政府定义(其中入门级和1级最低,8级最高)。
- 教育、[技能和培训]剥夺:与当地成年人、儿童和年轻人缺乏成就和技能相关的剥夺指数。基本数据包括教育成就水平、语言能力指标和儿童福利申请。
- 健康剥夺和残疾:基于过早死亡风险和因身体或精神健康不良而受损的剥夺指数。ONS用于计算这一指标的数据包括住院人数、死亡率和发病率。

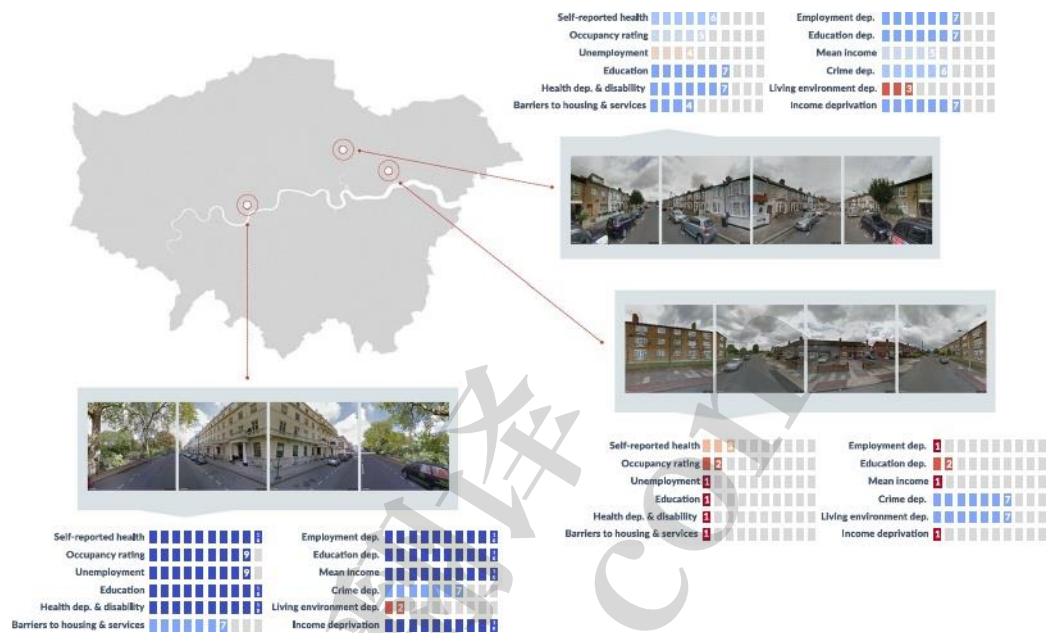


图 1 所示。分析中使用的街道图像和结果(标签)数据概述。通过指定相机方向以覆盖 360° 视角,为每个邮政编码获得四张图像。对于每个邮政编码,关于不同福祉结局的结果标签数据来自三个公共数据集:人口普查(Census)、英语剥夺指数(English Indices of Deprivation)和大伦敦当局(Greater London Authority)家庭收入估计。在分析中总共使用了 12 种不同的结果(见图 2 的结果及其定义方法)。对于每个结果,十分位数 1(红色)对应 LSOAs 中最糟糕的 10%,十分位数 10(蓝色)对应伦敦最富裕的 10%。结果标签被分配给 LSOA 内的所有邮政编码。图片由谷歌 Maps 提供。

- 自我报告的健康状况:报告自己的健康状况很差或不好的人的百分比;其他类别分别为尚可、良好和非常好。
- 就业剥夺:基于因失业、疾病或残疾或照顾责任而非自愿被排除在劳动力市场之外的成年人口的剥夺指数。基础数据来自政府支持索赔人。
- 失业:根据前一周的活动,推荐人(根据经济活动的优先顺序确定:全职工作、兼职工作、失业、退休、其他)被记录为失业的家庭的百分比。不从事经济活动的人,不积极寻找工作的人,退休的人,学生,照顾家人或家庭的人,或患有长期疾病或残疾的人不被归类为失业。
- 犯罪剥夺:与个人和物质受害风险有关的剥夺指数,包括暴力、入室盗窃、盗窃和基于警方记录的刑事损害。

我们从谷歌街景 API 获得了超过 100 万张图像,对应英格兰四大城市地区的 372,371 个邮政编码:大①伦敦②区、西米德兰兹区(包括伯明翰市)、③大曼彻斯特区和西约克郡(包括利兹市)。我们训练了不同的网络(在方法中详细描述),利用补充信息图 S1 中的架构,对每个邮政编码的四幅剪切图像进行有序分类,预测每个结局。为了评估预测的性能,我们使用了五折交叉验证,将已知但保留了十分位数的 LSOAs 的预测十分位数与实际十分位数进行比较(详细描述见方法)。

结果

伦敦福祉指标的空间模式和不平等。从政府统计数据中观察到的福祉结果数据显示,伦敦收入最高的地区在市中心和西南部,在伦敦市、肯辛顿、切尔西和威斯敏斯特的行政区(图 2A)。尽管伦敦的东区正在进行士绅化,并包含了一些富裕地区,但穷人大多生活在东部、东北部和东南部,以及西郊。和其他许多特大城市一样,过度拥挤也是一个问题,尤其是在市中心,这里空间有限,成本高昂,与收入无关⁷。居住环境的剥夺(住房质量、空气污染和道路安全的综合衡量)在内城更严重,甚至在较富裕的地区,并向郊区改善⁷。福祉结局还有其他复杂的空间分布,见图 2A 和补充信息图 S3。例如,收入剥夺与就业剥夺相关($r = 0.95$),但与住房和服务障碍只有部分重叠($r = 0.68$);生活环境中的剥夺和犯罪与所有其他指标的相关性较弱(平均 $r = 0.28$ 和 $r = 0.42$)。



图 2。训练有素的网络在预测伦敦福祉结局方面的表现:(A)观察到的和预测的平均收入、生活环境剥夺和犯罪剥夺的地图。(B)所有措施的观察和预测十分位数的比较。在(B)中，每个图显示了每个结局的预测和观察十分位数类别的数量，圆的大小代表了特定分配性能对应的 LSOAs 数量。完美分配只会对应于对角线上的大圆点，而随机分配则会等量均匀地覆盖整个分配空间。情节越接近前者，表现越好。

对幸福结果的预测。在所有结局中，未用于网络训练的 LSOAs 的真实十分位数和预测十分位数之间的平均 Pearson 相关性为 0.77;平均绝对误差(MAE)为 1.42。居住环境剥夺和平均收入预测效果最好

	分配精度			MAE	τ	r	k
	± 0	± 1	± 2				
结果变量							
自我报告健康	0.223	0.534	0.738	1.76	0.52	0.66	0.47
入住率评级	0.307	0.666	0.858	1.23	0.68	0.82	0.63
失业	0.255	0.590	0.800	1.48	0.61	0.76	0.55
教育	0.305	0.668	0.851	1.28	0.67	0.80	0.61
健康剥夺和残疾	0.256	0.610	0.820	1.41	0.63	0.79	0.57
住房和服务的障碍	0.233	0.562	0.781	1.59	0.58	0.73	0.52
就业不足	0.256	0.617	0.827	1.43	0.62	0.77	0.57
教育不足	0.276	0.640	0.842	1.33	0.65	0.80	0.60
平均收入	0.327	0.717	0.898	1.10	0.72	0.86	0.67
收入不足	0.281	0.653	0.854	1.30	0.66	0.81	0.61
犯罪被剥夺	0.195	0.466	0.673	2.02	0.44	0.57	0.39
生活环境剥夺	0.325	0.708	0.895	1.11	0.72	0.86	0.66

表 1. 多种福祉结局的预测结果:伦敦的分配性能。预测性能使用多个指标进行评估，每一列对应于不同的测量方法。这些措施包括测量正确预测类别的百分比的准确性指标，以及那些测量正确预测类别的百分比的允许误差范围分别为 ± 1 和 ± 2 类。我们还包括平均绝对误差(MAE)、肯德尔 τ 等级相关系数(τ)、皮尔逊相关系数(r)和使用线性权重(k)的科恩 $kappa$ 系数。

($r = 0.86$, $p < 0.01$;MAE 分别为 1.11 和 1.10)，其次是衡量过度拥挤程度的入住率($r = 0.82$, $p < 0.01$;MAE= 1.23)。表现最弱的是犯罪剥夺($r = 0.57$, $p < 0.01$;MAE= 2.02)和自我报告健康状况($r = 0.66$, $p < 0.01$;MAE= 1.76)(表 1)。

当考虑整个十分位数范围时，平均 62%(最小= 47%，最大= 72%)的 LSOAs 被分类在正确类别的 ± 1 以内，82%(最小= 67%，最大= 90%)被分类在 ± 2 以内(图 2B)。在区分代表极端不平等的最富裕和最贫穷的 LSOAs 方面，最贫穷和最贫穷地区的平均错误分类率分别为 4.6%和 3.4%(图 3)。在生活环境剥夺、平均收入和入住率方面，预测的最富裕和最贫穷地区的分布非常不同，只有 1.5%、1.9%、1.7%的最贫穷地区和 0.6%、0.4%、0.2%的最富裕地区被错误分类。相比之下，在犯罪剥夺方面，预测的最佳和最差分布存在部分重叠(分别有 12.2%和 11.6%的错分类)。这些结果表明，正如预期的那样，错误分类往往发生在相邻的十分位数之间，这些十分位数更难区分(例如，从十分位数 3 和十分位数 4 中区分图像)，但网络在区分最差和最好的区域方面表现良好。

学习的可移植性。为了评估学习的可转移性，我们从西米德兰兹的 53,577 个邮政编码获得了 199,576 幅图像，从大曼彻斯特的 66,029 个邮政编码获得了 246,056 幅图像，从西约克郡的 56,277 个邮政编码获得了 212,476 幅图像。当使用对伦敦数据进行训练的网络在其他城市进行预测时，在西米德兰兹、大曼彻斯特和西约克郡，所有指标的真实和预测 LSOA 类别之间的平均皮尔逊相关性分别为 0.68、0.71 和 0.66，平均 MAE 值分别为 1.72、1.59 和 1.75。与伦敦 LSOAs 相比，这些相关性较低，误差较大。然而，当我们使用目标城市的少量数据子集(使用每个城市所有可用数据的 1%、5%、10%和 20%)微调从伦敦获得的网络权重时，性能显著改善，并且与仅使用目标城市数据从头训练的结果相似。例如，当仅使用 1%的目标城市数据时(对应每个城市大约 500 个邮政编码位置)，平均相关系数增加到 0.78、0.74、0.74，MAE 下降到 1.38、1.49、1.49(图 4)。

优势与局限。我们研究的一个关键优势是，我们应用了相同的深度学习方法，在相同的空间尺度上检测多个结果之间的不平等，并使用同一组图像。因此，我们可以测试和比较用于评估是否以及对于哪些结果，可以使用街道图像来衡量城市不平等的表现。其次，通过测量图像的结果，我们的方法消除了选择和提取预定义特征的中间步骤。我们也不需要额外的数据集来将这些特征与感兴趣的结果联系起来。第三，对其他城市可转移性的评估可以帮助了解街道图像是否有可能在数据贫乏的地理位置作为低成本的监控工具，但不可避免地需要在未来的工作中进一步调查，具体如下所述。

我们研究的一个局限性是，用于训练网络的数据仅适用于那些有政府统计数据的措施，并且具有特定的空间分辨率。尽管如此，Isoa 级别的分辨率(平均人口 1614 人)高于大多数规划活动的分辨率。其次，目前街道图像仅在特定时间可用，这可能与用于模型训练和测试的数据收集时有所不同。在我们的分析中，图像和结果数据相差了几年(详见方法)。然而，模型性能很高，这表明图像包含了测量结果的视觉特征。及时和定期的数据集，特别是在快速变化的城市中，可以克服这一限制。第三，由于季节差异、天气条件和障碍物的影响，图像特征可能会有所不同。第四，

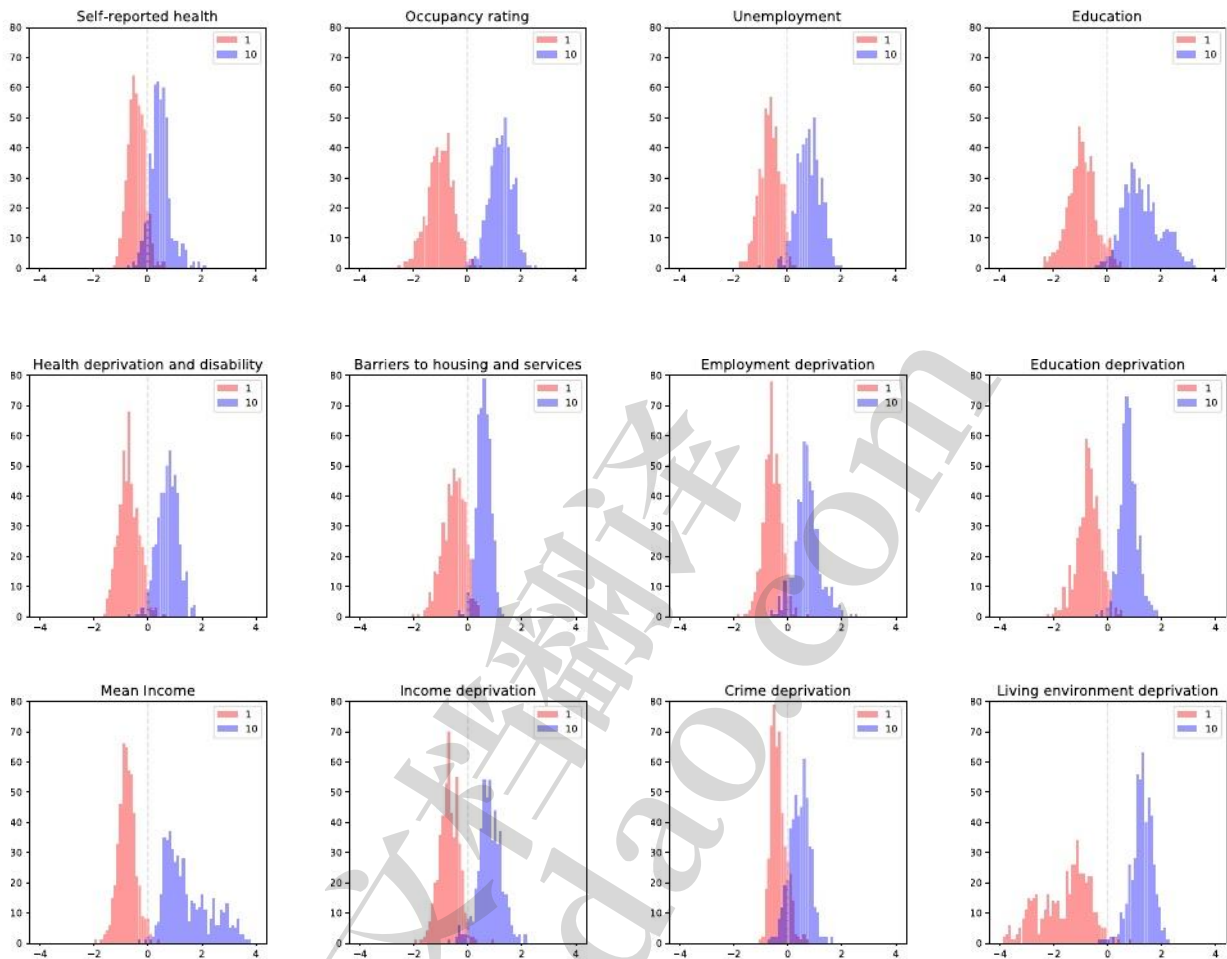


图 3. 训练有素的网络在区分伦敦最富裕和最贫穷地区方面的表现。每个单元格显示了在转换为十分位数类别之前，由训练网络的最后一层生成的每个 LSOA 的平均连续输出值的分布。根据地面真实数据，十分位数 1(红色)对应伦敦最糟糕的部分，十分位数 10(蓝色)对应最好的部分。这些极端十分位数的预测分布重叠越小，网络在区分它们(即，测量不平等的全部程度)。因此，与犯罪剥夺、自我报告的健康状况以及住房和服务障碍相比，该网络可以更容易地区分居生活环境剥夺、平均收入和占用率方面的最高和最低十分位数。

目前，网络的架构设计大多基于直觉、经验和实证评估⁴⁶，这是深度学习中的一个常见问题，因为目前在计算上搜索所有可能的架构是不可行的。集合模型等方法，用于统计模型的天气和健康预测集合统计模型^{47,48}，如果计算效率得到提高，也可以与深度学习网络一起使用⁴⁹。

可扩展性和未来研究。 几个未来的研究方向来自于我们的结果以及前面提到的数据和方法上的限制。首先，将这项工作扩展到其他城市和其他成果是理想的。这样做需要访问街道图像，目前可以通过私人数据提供商获得，但访问限制或成本限制了扩展到许多城市。众包图像或公共部门和政府在收集和提供图像方面的努力可以增加研究以及政策和规划目的的机会。其次，理想的做法是超越街头图像，在单一框架中使用多种数据输入，包括来自卫星、社交媒体和手机的数据输入，利用它们的互补信息来改进预测。要做到这一点，既需要数据访问，也需要开发适当而有效地使用多个数据源的方法。第三，应该进一步研究学习网络对其他具有相似或不同视觉特征(例如来自不同国家、发展轨迹、文化和气候)的地理位置的可移植性。如果可移植性是可行的，那么对于在数据丰富的环境中训练网络，并将其应用到数据有限的地方，将是有价值的。调查可移植性及其特征和决定因素，将需要不同环境下的结果数据;增加可移植性需要开发领域适应技术，以克服视觉特征的差异。第四，可移植性也应该随着时间的推移进行调查，在训练后的几年内使用训练过的网络进行测量，从而减少应该收集的新数据量。

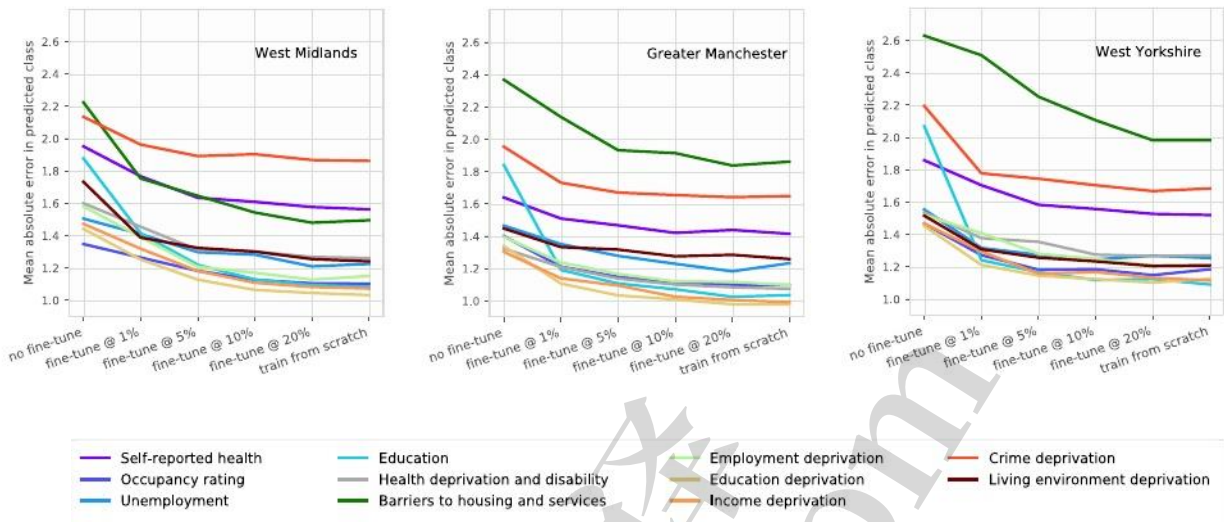


图 4. 该网络的性能训练伦敦数据应用于其他城市的图像。图表显示了在伦敦训练的网络的平均绝对误差，并使用其他三个目标城市的数据子集微调了网络权重。横轴显示了目标城市数据的使用程度，从直接使用伦敦训练好的网络权重在其他城市进行预测(没有微调)，到使用目标城市数据的子集微调网络权重(微调@ 1%、微调@ 5%等)，再到只使用目标城市数据而不转移权重的从头训练(从头训练)。即使只使用 1% 的目标城市数据，分配性能也会大幅提高，与从零开始训练时所获得的性能相似。虽然每个城市的不同结果在表现上存在差异，但表现最好和最差的结果组大致一致。

讨论和结论

在对多种社会、经济、环境和健康结果的比较、正面分析中，我们发现，将深度学习应用于街道图像，比其他结果(即犯罪、自我报告的健康)更好地预测了某些结果(即收入、生活环境)的不平等。

街道图像可以识别生活环境的变化，反映了一种直觉，即住房质量、污染(来源)和安全方面的道路状况等特征与视觉元素有关，就像占用率评级在一定程度上与人口密度有关一样。收入分配的表现也很高，因为低收入和高收入可能在住房、商店和车辆方面有视觉上的相关性⁴²。相比之下，观察到的较弱的犯罪表现与最近挑战所谓“破窗理论”的证据是一致的——破窗、涂鸦和垃圾堆积等视觉标志可能会导致犯罪的感知，但与更高的犯罪率无关⁵⁰。犯罪感知和犯罪率之间的区别可能解释了当使用众包感知安全指标时，街道图像的更高表现²⁴但对于实际犯罪率来说却不是这样。虽然自我报告健康的表现较低，但在健康剥夺的客观测量中观察到更高的准确性，该测量结合了入院率、死亡率和发病率的数据($r = 0.79, p < 0.01, MAE = 1.41$)。众所周知，自我报告的健康受到基于社会经济属性的健康预期差异的影响⁵¹。

我们还注意到，在一个城市的训练可以转移到同一国家其他城市的预测，特别是当网络仅用 1% 的目标城市图像进行微调时。精细调优支持的可转移性，如果在其他城市和其他国家复制，可能表明处于被剥夺和特权位置的共享视觉特征。这也意味着，通过训练有素的网络库，再加上可以有效收集的少量本地数据，可以实现高效的数据驱动的城市监控。

经济活动、创新、医疗保健、教育和其他公共服务的集中为城市提供了改善居民健康和福祉的潜力。然而，这些好处的分配却越来越不平等。许多城市意识到不平等加剧对社会不稳定和集体安全丧失的后果，开始利用其资源和权力来对抗不平等^{13,14}。为了解和衡量城市问责政策的影响，需要可靠和及时的数据，而传统来源的数据频率低或分辨率低，目前限制了这一点。图像有可能补充传统的基于调查和行政来源，以实现更高的时间和空间分辨率的监测。街道级图像目前主要是私营部门的领域(例如，谷歌，百度)，来自 Mapillary 等众包平台的贡献越来越大。这些数据虽然有价值，但存在不一致性，并且越来越多地受到访问限制。城市有能力要求或激励人们更好地获取这些数据，并利用自己的资产(例如多次穿越城市大部分地区的公交车和其他车队)收集数据，以增强我们衡量不平等现象的能力，并监测旨在解决不平等问题的政策的影响。

方法

数据源。图像数据。我们使用谷歌街景应用程序编程接口 (API) (<https://developers.google.com/maps/documentation/streetview/intro>) 获得了伦敦和英格兰接下来的三个最大城市(西米德兰兹郡包括伯明翰、大曼彻斯特和西约克郡包括利兹)的图像。我们首先获得了英国国家统计局(ONS)的邮政编码目录(<https://ons.maps.arcgis.com/home/item.html?id=1e4a246b91c34178a55aab047413f29b>), 并选择了分配给大伦敦行政区域 33 个地方当局区的 181,150 个邮政编码, 以及分配给其他三个城市的 175,883 个邮政编码。对于每个邮政编码, 如果可用, API 返回谷歌最近拍摄的最近可用全景图的唯一标识符('panoid');时间戳从 2008 年到 2017 年不等。

伦敦有 156 581 个邮政编码的全景图像, 相当于 131 465 个独特的全景图像。我们通过指定相机方向(即., 0° , 90° , 180° , 270°)相对于街景车辆覆盖 360° 的视野。因此, 我们总共获得了 525,860 张图像, 对应伦敦的 156,581 个邮政编码。对于其他城市, 我们从西米德兰兹的 53,577 个邮政编码获得了 199,576 幅图像, 从大曼彻斯特的 66,029 个邮政编码获得了 246,056 幅图像, 从西约克郡的 56,277 个邮政编码获得了 212,476 幅图像。

结果(标签)数据。我们从三个公开来源获得了关于人类福祉的多个维度的数据, 包括收入、健康、教育、就业、犯罪、住房和生活环境:2011 年人口普查(<https://www.ons.gov.uk/census/2011census>)、2015 年英语剥夺指数(<https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015>)和 2015 年大伦敦当局家庭收入估计(<https://www.data.london.gov.uk/dataset/household-income-estimates-small-areas>)。所有选定的结局数据均可在 LSOA 水平获得, 并从伦敦的 4,838 个 LSOAs、西米德兰兹的 1,680 个 LSOAs、大曼彻斯特的 1,673 个 LSOAs 和西约克郡的 1,388 个 LSOAs 获得。

对于每个结局, 我们计算了 LSOA 的十分位数, 十分位数 1 对应伦敦最糟糕的 10%, 十分位数 10 对应最富裕的 10%, 在可转移性分析中, 每个城市分别计算了十分位数。从邮政编码地点收集的个体图像与 LSOA 级别信息相匹配, 作为结果标签。

使用预训练卷积神经网络(convolutional neural network, CNN)进行特征提取。我们使用预训练的 CNN 作为固定特征提取器将 RGB 图像转换为 4096 维编码。我们使用 VGG16 网络⁵²使用来自 ImageNet 的 130 多万张图像进行训练⁵³。将街景图像及其预训练的权重作为该网络的输入, 并为每幅图像提取 fc6 层的 4096-D 输出。因此, 在这一步, 每个位置的特征特征由四个 4096-D 向量组成, 对应于每个位置的四幅图像。迁移学习, 使用在大型数据集上预训练的 cnn 作为固定的特征提取器, 已被证明有利于从头开始训练, 特别是在目标任务可以访问较少标记样本的场景中⁵⁴。

基于深度学习的十分位数分配。将每个邮编分配到一个结果十分位数是一个序数分类任务, 为此我们使用了补充信息图 S1 中所示的网络。如上所述, 我们使用了 VGG16 的预训练权重, 并且只训练了全连接层的权重。在这种架构中, 网络在架构中显示的四个通道中联合使用了来自每个位置的所有四张图像。来自不同通道的信息被聚合并馈入最后一层, 其中通过应用 sigmoid 函数计算 0 到 1 之间的单个连续值 p。批正归一化⁵⁵在网络中的所有层中使用, 除了在输出中计算 p。

为了解释十分位数之间的序数关系, 我们使用了 da Costa 和 Cardoso 提出的方法⁵⁶, 其中单个输出 p 值被解释为进行伯努利试验的概率, 即投掷一枚硬币, 得到正面的概率为 p, 反面的概率为 1-p, 正面和反面分别对应属于和不属于输出类的事件。对于 10 个(序数)输出十分位数类, 考虑 10 次不同的抛硬币, 并计算 $M = 1, \dots, 10$ 时获得 M 个正面的相应概率。这种方法将序数分类问题转换为具有 10 个类的分类任务。为了训练神经网络, 我们优化了以下交叉熵代价

$$\min_w \sum_n \sum_m y_n^m \ln p_n^m$$

其中 w 是网络权重, y_n 是 n 的标签向量吗? true 标签类的样本值为 1, 其他所有的样本值为 0。PNM 是 m 的概率⁵⁸十分位数为 n⁵⁸ 样本。在培训期间用于计算成本函数的真正标签是与分配邮编的 LSOA 相关的十分位数类。网络使用 Python 中的 TensorFlow 进行训练。我们使用 Adam 优化器⁵⁷学习率为 5e-6, 并训练网络进行了 10 万次迭代。在最后 5000 次迭代中产生最佳验证误差的网络被保留为最终模型。我们在补充信息中显示了训练集和验证集与训练周期相关的损失函数和平均绝对误差(为了简洁起见, 我们只显示了平均收入的结果;其他结果的行为是类似的)。

预测性能的测量。我们使用了五重交叉验证。在每次折叠中, 80%的数据(即., 80%邮编的图像-结果对)用于训练网络, 其余 20%被保留。然后, 我们测量了训练过的网络使用图像来预测训练中没有使用过的位置的结果的程度。我们重复了这个过程五次, 每次都拿出不同的 20%的数据。对于预测, 在应用前面提到的 sigmoid 之前, 通过平均最后一层训练网络产生的邮政编码级别的连续输出来计算每个 LSOA 的平均连续输出值

函数。然后将每个 LSOA 计算的平均值转换为十分位数类别，并与 LSOA 的实际十分位数进行比较。

我们使用 Pearson 的相关系数(r)、Kendall 的 tau 系数(τ)、Cohen 的 kappa (k)和平均绝对误差(MAE)，以及测量正确预测类的百分比的精度指标，以及分别测量允许误差幅度为 ± 1 和 ± 2 类的正确预测类的百分比的精度指标来评估性能。完整的结果如表 1 所示。

训练有素的网络可向伦敦以外的城市转移。我们评估了基于伦敦数据训练的模型在使用其他城市的图像预测感兴趣的结果方面的表现。首先，我们将来自伦敦的训练好的网络权重直接应用于来自每个目标城市的图像。其次，我们使用每个城市的邮编级别数据子集对网络权重进行微调，具体使用 1%、5%、10%和 20%的数据对网络进行微调。第三，我们只使用目标城市数据从头开始为每个城市训练网络。在正文中提供了每个城市的评价指标的总结。

数据可用性

本文中使用的数据集都是公开的，url 在“数据源”部分提供。

代码的可用性

发布后，代码将在 <https://github.com/esrasuel/measuring-inequalities-sview> 和 <http://equitablehealthycities.org/data-download/>上提供。

参考文献

1.联合国。《世界城市化展望:2018 年修订版》。《世界城市化展望》(2018 年)。

2.Ravallion, M., Chen, S. & Sangraula, P.全球贫困城市化的新证据。 *Popul. Dev. Rev.* 33, 667-701(2007)。

3.A.不平等、城乡差距和移民*。 *Q. J.经济学(Econ)*, 128, 1727-1785(2013)。

4.里昂,地区检察官,城市,城市化和健康。 *Int. J. epidemiology* . 37, 4-8(2008)。

5.《健康差距:不平等世界的挑战》, M.马莫特 著。(布卢姆斯伯里出版社, 2015)。

6.世界卫生组织和联合国人居署。《全球城市卫生报告:公平、更健康的城市促进可持续发展》(2016 年)。

7.英国国家统计局。2015 年英国贫困指数统计发布。(2015)。

8.Bennett, J. E.等。英格兰和威尔士预期寿命和预期寿命不平等的未来:贝叶斯时空预测。 *柳叶刀* 386,163-170(2015)。

9.城市会拉大贫富差距吗? *世界经济论坛*(2014)。

10.富有还是堕落?肯辛顿-切尔西:一个富裕但分裂严重的行政区。 *《经济学人》*(2017)。

11.刘志刚, 刘志刚。不平等与城市规模。 *启经济学.统计* 95,1535-1548(2013)。

12.联合国大会。改变我们的世界:2030 年可持续发展议程。大会第七十届会议(2015 年)。

13.大伦敦当局。 *增进所有伦敦人的健康——关于伦敦卫生不平等战略的磋商*。(2017)。

14.大伦敦当局。 *包容性伦敦:市长的平等、多样性和包容性战略*。(2018)。

15.瑞士联邦统计局。瑞士全国人口普查。查阅网址:<https://www.bfs.admin.ch/bfs/en/home/statistics/population/surveys/volkszaehlung.html>。(访问时间:2018 年 1 月 6 日)。

16.新西兰统计部长办公室。 *普查转型:前景光明*。(2015)。

17.瑞典研究理事会。Registerforskning.se。可在:www.registerforskning.se/en/获取。(访问时间:2019 年 1 月 29 日)。

18.Atkinson, J. & Blakely, T.新西兰的综合数据基础设施(IDI):迄今的价值和未来的机会。 *Int. J. Popul. 数据科学* 1(2017)。

19.van doorsler, E. & Gertham, u . g .自我健康评估的不平等是否预示着收入对生存的平等?来自瑞典数据的证据。 *Soc. 科学. 医学* 57,1621-1629(2003)。

20.Jean 等人。结合卫星图像和机器学习来预测贫困。 *Science* 353, 790-4(2016)。

21.Steele, J. E.等。利用手机和卫星数据绘制贫困地图。 *J. R. Soc. 界面* 14,20160690(2017)。

22.卫星图像可以精确地定位贫困,而调查却做不到。 *《纽约时报》*(2016)。

23.谢 M., Jean, N., Burke, M., Lobell, D. & Emon, S. S.基于深度特征迁移学习的遥感与贫困制图方法。发表于 AAAI 人工智能会议 (AAAI, 2016)。

24.Naik, N., Kominers, S.D., Raskar, R., Glaeser, E. L. & Hidalgo, C. A.计算机视觉揭示了物理城市变化的预测因素。 *Pnas* 114, 7571-7576(2017)。

25.萨里斯(P., Schechtner, K. & Hidalgo, C. A.)城市的协作形象:映射城市感知的不平等。 *PLoS One* 8, e68400(2013)。

26.Naik, N., philipom, J., Raskar, R. & Hidalgo, C. Streetscore-predicting the perceived safety of one million streetscape. IEEE 计算机学会计算机视觉和模式识别研讨会 793-799(2014)。

27.Dubey, a ., Naik, N., Parikh, D., Raskar, R. & Hidalgo, C. a .深度学习城市:在全球尺度上量化城市感知。载于 *《计算机视觉- ECCV 2016》*, 192-212(2016)。

28.Rzotkiewicz, a ., Pearson, A. L., Dougherty, B. V., Shontridge, A. & Wilson, N.系统评价谷歌街景在健康研究中的应用:主要主题、优点、缺点和未来研究的可能性。《健康场所》52,240-246(2018)。

29.Weichenthal, S., Hatzopoulou, M. & Brauer, M. A picture told thousand...exposure:在曝光科学和环境流行病学中深度学习图像分析的机遇和挑战。环绕。Int. 122, 3-10(2019)。

30.J. V. Henderson, Storeygard, A. & Weil, D. N.从外太空测量经济增长。《美国经济评论》102,994-1028(2012)。

31.Mathieu, R., Freeman, C. & Aryal, J.使用面向对象技术和非常高分辨率的卫星图像绘制城市地区的私人花园。Landsc. 城市的计划。(2007)。

32.Weeks, J.R., Hill, a ., Stow, D., Getis, a . & Fugate, D.我们能从空中发现一个社区吗?加纳阿克拉的社区结构定义。GeoJournal 69, 9-22(2007)。

33.伦敦人民的生活和劳动。(1902)。

34.桑普森, R. J. & Raudenbush, S. W.看失序:社区污名与“破窗”的社会建构。Soc. Psychol. Q. 67, 319-342(2004)。

35.桑普森, R. J. & Raudenbush, S. W.公共空间的系统社会观察:城市社区无序的新视角。点。社会科学。105,603-651(1999)。

36.ogers, C. L., Caspi, a ., Bates, C.J., 桑普森, R. J. & Moffitt, T. E.基于谷歌街景的儿童社区系统社会观察:一种可靠且成本效益高的方法。J.儿童心理学。《精神病学》53,1009 - 1017(2012)。

37. Morland, K., Wing, S., Diez Roux, A. & Poole, C. 与食品商店和食品服务场所的位置相关的社区特征。 *点。 J. Prev. Med.* 22, 23–9 (2002)。
38. Kuffer, M., Pfeffer, K. & Sliuzas, R. 贫民窟来自空间利用遥感绘制的 15 年贫民窟地图。 *《遥感8》* (2016)。
39. 罗斯, 社区劣势和成人抑郁症。 *《健康与社会行为杂志》* 44(2), 177–187 (2000)。
40. “破窗”与自尊: 社区贫困与混乱的主观理解。 *Soc. 科学.* 第 36 号, 968–994 (2007)。
41. 科恩, D.A., 法利, T. A. & 梅森, K., 为什么贫穷不健康? 社会和物质的中介。 *Soc. 科学. 医学* 57, 1631–1641 (2003)。
42. Gebru 等。使用深度学习和谷歌街景来估计美国各地社区的人口构成。 *Pnas* 114, 13108–13113 (2017)。
43. Arietta, S.M., Efros, A.A., Ramamoorthi, R. & Agrawala, M. 城市取证: 使用视觉元素预测非视觉城市属性。 *IEEE 反式. 粘度第一版. 图.* 20, 2624–2633 (2014)。
44. LeCun, Y., Bengio, Y. & Hinton, G. 深度学习。 *Nature* 521, 436–444 (2015)。
45. UK 国家统计局, 2011 年人口普查。 查阅网址: <https://www.ons.gov.uk/census/2011census>。(访问时间: 2019 年 1 月 24 日)。
46. Goodfellow, I., Bengio, Y. & Courville, A. 深度学习。(MIT 出版社 2016)。
47. 用集合方法进行天气预报。 *科学* 310, 248–9 (2005)。
48. Kontis, V. 等人。35 个工业化国家的未来预期寿命: 贝叶斯模型集合的预测。 *柳叶刀* 389, 1323–1335 (2017)。
49. Krogh, A. & Vedelsby, J. 神经网络集成, 交叉验证和主动学习。《神经信息处理系统的进展》第 7 期 (1995)。
50. O'Brien, D. T. & Sampson, R. J. 社区混乱的公共和私人领域。 *《刑事犯罪》* 52, 486–510 (2015)。
51. Salomon, J.A., Tandon, A. & Murray, C. J. L. 自评健康的可比性: 使用锚定小插曲的横断面多国调查。 *BMJ* 328, 258 (2004)。
52. Simonyan, K., Vedaldi, A. & Zisserman, A. 深层卷积网络: 可视化图像分类模型和显著图。预印本在 <https://arxiv.org/abs/1312.6034> (2013)。
53. Russakovsky, O. 等人。ImageNet 大规模视觉识别挑战赛。 *Int. J. 康普特. Vis.* 115, 211–252 (2015)。
54. Yosinski, J., Clune, J., Bengio, Y. & Lipson, H. 深度学习中的特征有多可转移? *ad. Neural Inf. Process. Syst. 27 (Proceedings NIPS)* 27, 1–9 (2014)。
55. 批归一化: 通过减少内部协变量移位来加速深度网络训练。预印本在 <https://arxiv.org/abs/1502.03167> (2015)。
56. da Costa, J. P. & Cardoso, J. S. 使用神经网络的有序数据分类。in (eds Gama, J., Camacho, R., Brazdil, P. B. & Jorge, a.m.t. L.) 690–697 (施普林格, 柏林, 海德堡, 2005)。
57. 王志强, 王志强, 王志强: 一种随机优化方法。预印本: <https://arxiv.org/abs/1412.6980> (2014)。

确认

本研究得到了惠康信托基金(赠款 205208/Z/16/Z 和 209376/Z/17/Z)和英国 HDR UKRI 创新/卢瑟福基金奖学金 (UKRI Innovation/Rutherford Fund Fellowship at HDR to ES) 的支持。该论文得益于与 Michael Brauer、Seth Flaxman、Ender Konukoglu 的讨论。

作者的贡献

E.S. 在 J.P. 和 M.E. 的输入下设计了研究概念在美国, E.S. 获取数据, 进行分析, 并在 J.E.B. 和 M.E. 的输入下准备结果 E.S. 和 M.E. 写了论文的初稿。J.P. 和 J.E.B. 为最后的论文做出了贡献。

额外的信息

本文的补充信息见 <https://doi.org/10.1038/s41598-019-42036-w>。竞争利益: 作者声明没有竞争利益。

出版者注: 施普林格 Nature 对已出版地图和机构附属关系中的管辖权要求保持中立。



开放获取 本文遵循创作共用属性 4.0 国际许可协议 (Creative Commons Attribution 4.0 International License) 授权, 该协议允许以任何媒介或方式使用、分享、改编、分发和复制

格式, 只要您对原作者和来源给予适当的信任, 提供知识共享许可的链接, 并说明是否做出了更改。本文中的图像或其他第三方材料包括在本文的知识共享许可中, 除非在材料的信用额度中另有说明。如果材料未包括在文章的知识共享许可中, 并且您的预期用途不被法定法规允许或超过允许的用途, 您将需要直接获得版权持有人的许可。如需查看本许可协议副本, 请访问 <http://creativecommons.org/licenses/by/4.0/>。

有道文档翻译
pdf.youdao.com