arXiv:2304.01404v1 [cs.LG] 3 Apr 2023

基于主动迁移学习水平集估计的材料表面自适应缺陷区域识别

ShotaHozumi[1, †], Kentaro Kutsukake[2, 3, †], KotaMatsui[4, *], Syunya Kusakawa[1], ToruUjihara[3,5], IchiroTakeuchi[6, 2, *]

[1] 日本名古屋市昭和区御器所町，名古屋工业大学计算机科学系，邮编466-8555[2] 日本东京都中央区日本桥1丁目1-4-1，理化学研究所先进智能项目中心，邮编103-0027[3] 日本名古屋市千种区不老町，名古屋大学材料与系统可持续性研究所，邮编464-8603。[4] 日本名古屋市昭和区鹤舞町65号，名古屋大学生物统计学系，邮编466-8550[5] 日本名古屋市千种区不老町，名古屋大学材料过程工程系，邮编464-8603[6] 日本名古屋市千种区不老町，名古屋大学机械系统工程系，邮编464-8603

[†] 这些作者对本研究的贡献相同。

关于材料的通信和索取事宜，请联系 K.M.（电子邮箱：matsui.k@med.nagoya- u.ac.jp）和 I.T.（电子邮箱：ichiro.takeuchi@mae.nagoya-u.ac.jp）。

摘要

在材料表征中，识别材料表面的缺陷区域至关重要。传统方法是在表面预先确定的网格点上逐点测量相关物理性质，并确定该性质未达到预期水平的区域。为更高效地识别缺陷区域，我们提出了自适应映射方法，优先利用测量资源检测缺陷区域的边界。我们将这一问题视为水平集估计（LSE）问题中的主动学习（AL）问题。基于AL的LSE旨在以尽可能少的测量次数，确定定义在表面上的物理性质函数的水平集。此外，为应对反复生产具有相似规格的材料的情况，我们引入了一种迁移学习方法，以便有效利用先前生产的材料所包含的信息。作为概念验证，我们将所提出的方法应用于硅晶圆的红区估计问题，并证明，与传统方法相比，我们能够以显著更低的测量成本识别缺陷区域。（179 个单词）

关键词

材料表征、空间映射、主动学习、贝叶斯建模、水平集估计、迁移学习

引言

评估材料表面相关物理性质的空间分布至关重要 [1, 2, 3, 4]。在制造业中，识别材料表面上物理性质未达到预期水平的缺陷区域尤为重要。例如，在半导体材料的制造过程中，必须识别硅砖表面上载流子寿命低于预期水平的缺陷区域 [5, 6, 7]。在传统的缺陷区域识别问题中，通常会在均匀分布于材料表面的预定网格点上逐点测量物理性质。遗憾的是，当物理性质测量成本高昂时，这种传统方法效率低下。为了提高效率，可以在明显属于正常区域或缺陷区域的地方减少测量点，而在正常区域与缺陷区域的边界附近增加测量点。然而，在确定材料表面的物理性质分布之前，无法采用这种自适应方法。因此，有必要开发一种方法，在估计物理性质分布的同时，实现测量点的自适应选择。

在本研究中，我们将缺陷区域识别问题视为水平集估计（LSE）问题 [8, 9, 10, 11, 12, 13]，：我们将材料表面上的物理性质分布看作定义在二维输入空间上的实值函数，并估计其水平集——即函数取给定常数值时对应的输入空间集合。当将缺陷区域定义为材料表面上物理性质低于某一阈值 θ 的点时，θ 处的水平集可解释为正常区域与缺陷区域之间的边界。在此，我们介绍基于主动学习（AL）的水平集估计方法。这类方法通过自适应选择待评估的点，以尽可能少的函数评估次数估计函数的水平集。主动学习是监督式机器学习（ML）的一种特殊形式，其学习算法能够自适应地选择训练集 [14, 15, 16, 17, 18]，中的输入点；在统计学文献中，这也称为最优实验设计 [19, 20, 21, 22]。基于AL的LSE方法能够自适应地选择材料表面上的测量点，同时有效搜索正常区域与缺陷区域的边界。

当反复生产规格相同的材料时，可以参考先前生产材料的物理性质空间分布，估计新生产材料的正常区域和缺陷区域。在机器学习（ML）中，

利用过去相似问题的解高效解决新问题的方法称为迁移学习（TL） [23]。本研究提出了一种用于水平集估计问题（LSE）的迁移学习方法，称为基于主动迁移学习（ATL）的 LSE。通过有效迁移先前生产材料的知识，基于 ATL 的 LSE 方法只需比不进行知识迁移的方法更少的测量，即可识别缺陷区域。在制造过程中，环境和操作条件的变化很容易导致位置和尺度偏移。在这种情况下，传统迁移学习方法并不足够，因为这类方法仅假设源（旧）问题与目标（新）问题的两个函数相似。为解决这一问题，我们提出了一种基于 ATL 的 LSE 方法变体，即使存在位置和／或尺度偏移也可应用。该变体利用位置和／或尺度发生变化后源问题与目标问题之间的相似性。

本研究的动机源于太阳能电池硅锭的红区估计问题 [6, 7, 24, 25, 26, 27]。硅锭是太阳能电池的基础材料，其质量指标之一是载流子寿命。载流子寿命是指受光激发的载流子返回基态所需的时间，与太阳能电池的转换效率呈正相关。在实际的硅锭生产过程中，杂质会进入生产时与坩埚接触的材料外部区域，导致载流子寿命值较低。材料表面载流子寿命值较低的区域称为红区。因此，在将材料送入下一制造阶段之前，识别硅砖表面的红区（缺陷区域）至关重要。作为概念验证，我们将所提出的基于主动学习和基于 ATL 的 LSE 方法应用于太阳能电池硅锭的红区估计问题，以展示其有效性。

结果

太阳能电池硅锭的红区估计 本研究的动机源于太阳能电池用多晶硅锭的红区估计问题，这类硅锭是太阳能电池基片的常用材料（图1a）。硅锭外围会受到坩埚壁杂质扩散的污染，因而无法用于太阳能电池。这些无法使用的区域称为红区，其分布和体积是影响硅锭材料成品率的重要参数 [24, 25, 26, 27]。传统上，红区被定义为载流子寿命低于阈值的区域，并通过空间映射测量加以识别。载流子寿命是半导体材料的一项基本物理性质，与太阳能电池性能呈正相关。在大多数情况下，测量会沿等间距网格坐标逐点进行（图1b）。然而，这种网格映射耗时较长，其成本已成为材料开发的瓶颈。本研究采用基于AL的LSE方法，减少载流子寿命测量所需的测量点数量（图1d）。与一般工业生产中的情况一样，不同硅锭的红区空间分布相似；但每件产品的载流子寿命值的位置和尺度往往有所偏移。为处理这些特征，我们将基于 ATL 的 LSE 方法应用于该问题（图1e）。我们研究的问题是：估计一块硅锭中硅砖表面的红区，该表面包含6,586个网格点。作为概念验证，我们预先测量所有网格点的载流子寿命值，并通过检验LSE方法能否仅凭少量选定点的测量值正确识别真实红区，来评估这些方法的性能。

问题设定：水平集估计（LSE）我们将识别缺陷区域的问题（例如太阳能电池铸锭中的红区）视为一个 LSE 问题。令 $f(x)$ 表示二维材料表面位置 x 处的相关物理性质，例如载流子寿命。如果要求该物理性质不低于某个阈值 θ，则可将缺陷区域识别问题表述为识别上水平集 $\mathcal{U}_\theta = \{x \in \mathcal{X} \mid f(x) \geq \theta\}$ 和次水平集$\mathcal{L}_\theta = \{x \in \mathcal{X} \mid f(x) < \theta\}$的问题，其中 $\mathcal{X}$ 是整个输入域，即整个材料表面。假设物理性质在二维材料表面上平滑变化，则可通过在足够密集的网格点 $x_1, \ldots, x_N \in \mathcal{X}$ 测量物理性质，来估计函数 f 在 θ 处的上水平集和次水平集，

其中 N 是网格点的数量。然而，如果测量成本很高，在全部 N 个网格点采集观测值的总成本可能高得难以接受。我们的基本思路是，只在部分特定网格点上评估函数 f，以降低总测量成本；在这些网格点进行函数评估，有助于估计 θ 处的上水平集和次水平集。也就是说，在物理性质明显高于或低于阈值 θ 的位置进行测量效率不高，而在物理性质接近 θ 的位置有效利用测量资源则更为合理。基于这一思路，我们提出一种用于缺陷区域识别问题的主动学习（AL）方法，旨在尽可能减少函数评估次数来估计 θ 处的水平集。在以下各节中，我们介绍基于 AL 和基于主动迁移学习（ATL）的方法，通过自适应地选择测量点来有效解决水平集估计问题。图2展示了一个水平集估计问题的示例。

基于主动学习（AL）的水平集估计（LSE）方法 AL 是一种监督式机器学习方法，可选择训练输入点，从而以尽可能少的函数评估来估计函数。针对水平集估计问题，已有多种基于 AL 的方法 [8, 9, 10, 11, 12, 13]。基于 AL 的 LSE 方法的基本思路是采用贝叶斯回归模型近似物理性质函数 f。基于 AL 的 LSE 方法通过将选定的测量数据加入训练集，迭代更新贝叶斯回归模型。本研究采用高斯过程（GP）模型 [28] 作为贝叶斯回归模型。假设当前处于步骤 t，此时已测量物理性质的少量网格点 $x_i,\ i \in \mathcal{S}_t \subset \{1,\ldots,N\}$ 其中 $\mathcal{S}_t \subset \{1,\ldots,N\}$ 是网格点索引的一个子集。随后，使用带标签样本集 $\{(x_i,\ f(x_i))\}_{i\in\mathcal{S}_t}$ 拟合贝叶斯回归模型。贝叶斯回归模型的优势在于，它不仅能提供每个网格点处的预测均值，还能提供预测方差。预测方差表示每个网格点处物理性质的不确定性。基于网格点处函数值的均值和方差，选择一个新点 $x_{i(t)},\ t \notin \mathcal{S}_t$，该点有望有效估计 θ处的超水平集和次水平集，并测量所选位置 $x_{i(t)}$ 处的物理性质 $f(x_{i(t)})$。接着，在下一步 $t+1$中，子集更新为 $\mathcal{S}_{t+1} \leftarrow \mathcal{S}_t \cup \{i(t)\}$，并重复相同过程。重复这些步骤，直至充分准确地估计出 θ 处的超水平集和次水平集。图3展示了一维函数的基于 AL 的 LSE 方法。

图4a 展示了基于 AL 的 LSE 方法某一步骤的快照。基于 AL 的 LSE 方法的核心思路是利用每个网格点处物理性质的可信区间；若可信区间下界大于阈值 θ，则物理性质很可能大于 θ；若可信区间上界小于阈值 θ，则物理性质很可能小于 θ。也就是说，应将测量资源有效用于可信区间包含阈值 θ 的网格点，因为这表明物理性质究竟大于还是小于阈值仍不明确。具体而言，下一个评估点根据违反度选取，其定义为 $\alpha(x) = \min\{\max\{0, \theta - \ell(x)\}, \max\{0, u(x) - \theta\}\}$。图4b 展示了若干输入候选点的违反度（详情见方法和补充信息1-3s）。

评估标准 根据可信区间，每个未测量点 x_i，$i \notin \mathcal{S}_t$可分为以下三类：

- $f(x_i) > \theta$（如果下限大于 θ）；
- $f(x_i) < \theta$（如果上限小于 θ）；
- 未确定（否则）。

这意味着，在每一步中，基于主动学习的水平集估计方法的优良度都是根据下表中每个单元格的计数进行评估的：

		基于主动学习的水平集估计方法的预测		
		$\hat{f}(x_i) > \theta$	未确定	$\hat{f}(x_i) < \theta$
真实值	$f(x_i) > \theta$	TP	UP	FN
	$f(x_i) < \theta$	FP	UN	TN

其中，TP、FN、FP 和 TN 分别表示真正例、假阴性、假正例和真负例，f 表示基于高斯过程模型的预测。此外，我们将基于主动学习的水平集估计方法的输出为“未确定”的情况分别称为未确定阳性（UP）和未确定阴性（UN）。如何处理这些未确定点取决于问题设定。如果去除缺陷区域比保留正常区域重要得多，则应将未确定点视为（可能的）缺陷区域。在这种情况下，所有未确定点都视为次水平集 $\mathcal{L}_\theta$的元素。特别是，如果 $\boldsymbol{x}$ 是一个未确定正例（UP 实例），则应将 $\boldsymbol{x}$ 视为假阴性（FN），并在灵敏度和特异度的计算中使用 UP + FN 作为

假阴性。类似地，如果 $\boldsymbol{x}$ 是一个未确定负例（UN 实例），则应将 $\boldsymbol{x}$ 视为真负例（TN），并在灵敏度和特异度的计算中使用 UN $+$ TN 作为真负例。这便得到以下“风险敏感型”灵敏度和特异度：

$$\begin{aligned}\text{sensitivity}_{\text{risk-sensitive}} &= \frac{\text{TP}}{\text{TP}+\text{UP}+\text{FN}},\\ \text{specificity}_{\text{risk-sensitive}} &= \frac{\text{TN}+\text{UN}}{\text{TN}+\text{FP}+\text{UN}}.\end{aligned}$$

另一方面，如果维持正常区域远比移除缺陷区域重要，则应将未确定点视为（可能的）正常区域。在这种情况下，所有未确定点都视为上水平集 $\mathcal{U}_\theta$中的元素。具体而言，如果 $\boldsymbol{x}$ 是一个未确定正例（UP instance），则 $\boldsymbol{x}$ 应视为真正例（TP），我们在计算灵敏度和特异度时将 UP $+$ TP 作为真正例。如果 $\boldsymbol{x}$ 是一个未确定负例（UN instance），则 $\boldsymbol{x}$ 应视为假正例（FP），我们在计算灵敏度和特异度时将 UN $+$ FP 作为假正例。由此得到以下“成本敏感型”灵敏度和特异度指标：

$$\begin{aligned}\text{sensitivity}_{\text{cost-sensitive}} &= \frac{\text{TP}+\text{UP}}{\text{TP}+\text{UP}+\text{FN}},\\ \text{specificity}_{\text{cost-sensitive}} &= \frac{\text{TN}}{\text{TN}+\text{FP}+\text{UN}}.\end{aligned}$$

利用这两类指标，我们同样定义两类曲线下面积（AUC）和 F_1 分数，它们是评估模型性能 [29]的标准指标。

基于AL的LSE方法的结果 我们将AL-LSE方法应用于太阳能电池铸锭的红区估计问题。此处，74×89 个网格点均匀分布在尺寸为186毫米 × $\times$ 156 毫米的硅锭表面上。我们将所提出的方法与两种基线方法进行了比较。第一种基线方法是随机方法，即均匀随机地选择下一个测量点。第二种基线方法是非自适应方法，其示意图见补充图1（实验设置详情见方法）。图5a、b和c 分别展示了随机方法、非自适应方法和基于AL的LSE方法的行为（补充视频1–3展示了各方法在每一步如何选择测量点）。在这些图中，蓝色、红色和灰色区域分别表示估计的超水平集、次水平集和未确定区域。图6展示了太阳能电池铸锭数据上随机方法（红色）、非自适应方法（紫色）和基于AL的LSE方法（蓝色）的风险敏感型AUC、风险敏感型F_1 分数、成本敏感型AUC和成本敏感型F_1 分数，以及使用迁移学习（TL）的方法的相应指标；下文将介绍后者。

用于太阳能电池铸锭数据的LSE（蓝色）方法，以及采用迁移学习（TL）的方法；下文将对此进行介绍。

通过比较**a**、**b**和**c**中的图可以看出，在早期阶段（见左列），三种方法之间没有明显差异；然而，随着步骤推进（见中列），基于主动学习的水平集估计方法在左右两端都有更多测量点，而这些位置正是实际的缺陷区域。此外，在最后阶段（见右列），基于主动学习的水平集估计方法在正常（蓝色）区域与有缺陷（红色）区域的边界附近布置了许多测量点。基于主动学习的水平集估计方法几乎将所有区域都判定为正常（蓝色）或有缺陷（红色），而随机方法和非自适应方法即使在最后阶段仍有一些灰色区域。图**6**清楚地显示了主动学习方法相较于两种基线方法的优势。请注意，在风险敏感和成本敏感情景下，两种基线方法的性能指标差异显著，因为仍存在未确定区域；这些区域无法被判定为正常或有缺陷。

主动迁移学习（ATL） 为了更高效地识别新生产材料的红区，可以利用先前生产材料的测量数据。我们使用迁移学习（TL）框架实现了这一思想，其中旧问题和新问题分别称为源域和目标域。TL涵盖了一大类在各种机器学习任务中利用额外信息的方法，包括特征迁移和模型迁移方法。在本研究中，我们采用了实例迁移 [23]方法，其中先前开展的实验（源域）中的数据被有效地用于当前实验~~（目标域）~~。~~针~~对实例迁移，已有多种方法被提出，例如重要性加权 [30] 和最优传输 [31], ，我们采用了Diff-GP方法 [32] ，因为它可以直接使用上文讨论的高斯过程模型。在Diff-GP中，目标函数与源函数之间的差异通过高斯过程回归模型进行建模。由于假设源问题与目标问题相似，因此只需少量训练实例即可高精度地估计差异函数。设 $\{(x'_j, y'_j)\}_{j=1}^m$ 为源问题的^训练实例。那么， $\{(x'_j, y'_j + f_{\text{diff}}(x'_j))\}_{j=1}^m$ 可以作为目标问题的^额外训练实例，其中 f_{diff} 为估计的均值差异函数。在所提出的基于 ATL 的 LSE 方法中，通过纳入差异函数的不确定性来选择测量点。图**7a-c** 说明了Diff-GP的基本概念。

尽管不同产品的物理性质空间分布通常相似

但由于测量条件和材料质量的波动，各产品的位置和尺度往往会发生偏移。为解决这一问题，我们还开发了一种基于 ATL 的 LSE 方法变体，即使位置和/或尺度发生偏移也适用。该变体利用位置和/或尺度变化后源域与目标域之间的相似性。当源域经过位置-尺度偏移后与目标域相似时，$\{(x'_j,\ \gamma(y'_j+f'(x_j))+\eta)\}_{j=1}^{m}$可作为目标域的额外训练实例，其中 η 和 γ 分别为位置和尺度偏移参数。在这一基于 ATL 的 LSE 方法变体中，测量点（针对目标域）的选择方式使差异函数和位置-尺度偏移参数能够同时估计。图 7**d–f** 展示了位置-尺度偏移的基本概念。

图 5 和图 6 展示了基于 ATL 的 LSE 方法以及基于位置-尺度偏移（LSS）-ATL 的 LSE 方法的结果。在图 5 中，与基于 AL 的 LSE 方法相比，基于 ATL 和 LSS-ATL 的 LSE 方法在早期阶段就能将许多区域识别为正常（蓝色）或有缺陷（红色）（见左列）。这是因为通过迁移学习有效利用了先前生产材料的信息。另一方面，在本次实验中，基于 ATL 和 LSS-ATL 的 LSE 方法在性能上没有显著差异。这是因为本次数据来自一项控制良好的实验，源硅锭与目标硅锭之间不存在显著的环境或操作变量。此外，图 6 表明，在全部四项性能指标上，利用迁移知识都具有明显优势。这些结果表明，将先前生产材料的知识纳入迁移学习框架是有效的。

讨论

本研究表明，与采用预设测量点的传统方法相比，自适应地选择测量点能够以低得多的测量成本，有效识别材料表面的缺陷区域。如图 5 所示，这一问题可以解释为将材料表面划分为正常区域和缺陷区域的问题。这意味着，选择测量点时应尽可能缩小未确定区域（图 5 中的灰色区域），即尚无法判定为正常或有缺陷的区域。本文的主要贡献是表明这一问题可以表述为水平集估计问题，并且主动学习方法能够为该问题自适应地选择有效的测量点。事实上，在硅锭红区识别问题中，与传统方法相比，我们能够以显著更低的成本识别缺陷区域。此外，我们证明，利用具有相似规格的先前生产材料中的信息这一框架，可以视为迁移学习问题，并据此提出了基于 ATL 和 LSS-ATL 的 LSE 方法。在制造业中，当材料表征的测量成本较高时，抽样检验十分常见。采用所提出的主动迁移学习方法，有望通过降低每种材料的检验成本，实现全面检验。

在此，作为一项以硅锭为对象的概念验证研究，基于 AL 和 ATL 的 LSE 方法显著减少了样品表面载流子寿命映射所需的测量点数量。这一减少得益于硅砖载流子寿命的特性：其分布连续，且不同产品之间相似。这些是工业生产材料的物理性质所共有的特征。因此，基于 AL 和 ATL 的 LSE 方法有助于评估各种材料的各种物理性质的空间分布。相比之下，这一优势也表明，这些方法难以应用于物理性质呈不连续且特定分布的材料。对于多晶硅锭而言，识别位错簇就是一个此类难题。除红区外，由位错簇造成的低寿命区域也是多晶硅锭中的缺陷区域，如图1c所示。位错簇是在凝固过程中由硅锭内部的应力产生的，其位置因硅锭而异。因此，若将先前生产的硅锭中位错簇位置的信息迁移到新生产硅锭的测量中，基于 ATL 的 LSE 方法的性能可能会下降。

为应对材料表面物理性质的不连续变化，可在高斯过程模型中采用适当的核函数，例如 Matérn核。Matérn核因能够捕捉不连续变化，已应用于空间统计学、地统计学和图像分析领域 [33, 28]。通常，应根据物理性质的特征选择适当的核函数和超参数。尽管在能够获取先前生产材料的信息时，我们通过交叉验证选择了超参数，但迁移超参数信息也很有效 [34]。在迁移学习中，将源域信息迁移到目标域可能会形成不恰当的偏差，这种情况称为负迁移 [23]。使用先前生产材料的信息时，我们需要谨慎判断是否可能发生负迁移。据我们所知，目前尚无在主动学习设置下预测负迁移的方法。不过，在实际应用中，我们可以引入启发式机制来做出自适应决策，例如，当目标域中的函数值与迁移信息预测的函数值存在显著差异时，不使用迁移学习。尽管如此，许多工程问题涉及反复生产规格相似的材料。在此类情况下，正如本研究所示，基于自适应迁移学习的方法是有效的。

材料表面的物理性质映射还涉及测量探头移动距离的问题。如图**1b**所示，在传统的网格映射中，物理性质测量时各点之间的移动距离最短。相比之下，在所提出的基于 AL 和 ATL 的方法中，下一个测量点从整个表面区域中选取。因此，探头的总移动距离会比传统方法更长。在某些情况下，尽管测量点数量减少，但由于探头移动耗时，总测量时间反而会增加。对于探头移动需要一定时间的映射设备而言，这是一个严重的问题。为了将探头移动成本纳入测量位置的自适应选择中，我们需要引入一种新的主动学习机制。许多主动学习方法，包括本文讨论的方法，都被称为短视方法，因为每一步选择下一个训练实例时，都没有考虑其对后续步骤的影响。要克服这一挑战，我们需要引入非短视主动学习 [35] 或强化学习 [36, 37], 的理念；后者已被广泛研究，可用于不确定环境下的最优控制。

方法

载流子寿命数据 本研究所用的载流子寿命数据取自一块 G2 尺寸的半工业级多晶硅锭（$400 \times 400 \times 200$ mm^3，重 70 kg）。每块硅砖的侧表面积为 156×186 mm^2, ，并以间隔 2 mm 排列的网格点作为测量位置，共计 6,586 个网格点。寿命值低于阈值 $\theta = 2.0$ 的网格点被定义为红区。为检验 LSE 方法能否正确识别红色区域，预先测量了所有网格点的寿命值。微波光电导衰减测量是在未钝化的硅砖切片表面上进行的。不同硅锭之间载流子寿命的位置-尺度偏移源于表面状况和材料质量的差异。在 LSE 方法的实验中，假设这些网格点的寿命值未知，并从中选取检测点；通过评估 LSE 方法将多少网格点正确分类为红区或正常，来衡量其性能。

问题表述 假设给定一个黑箱函数 $f : \mathcal{X} \to \mathbb{R}$ 和候选点$x_1, ..., x_N \in \mathcal{X}$。此处， $\mathcal{X}$表示材料表面上的一组输入位置，函数 f 的输出是一个实值相关物理性质。此外，输入 x 时函数值 $f(x)$的观测值 y 表示为 $y = f(x) + s$，其中含有均值为 0、方差为 σ^2的高斯误差 $s \sim \mathcal{N}(0, \sigma^2)$。LSE [9] 问题被表述为将每个输入点 x_i 划分到以下两个区域的问题

$$\mathcal{U}_\theta = \{x \in \mathcal{X} \mid f(x) \geq \theta\} \quad \text{or} \quad \mathcal{L}_\theta = \{x \in \mathcal{X} \mid f(x) < \theta\} \tag{1}$$

对于某个阈值 θ。因此，缺陷区域识别问题可视为一个输入为二维实值黑箱函数 f 的 LSE 问题。在常规回归分析中，可以通过利用我们已作为训练数据获得的多组输入和输出 $\mathcal{D}_N = \{(x_i, y_i)\}_{i=1}^N$ 来估计 f，从而对新的输入执行 (1)。然而，在材料科学等许多科学领域，观测输入 x 对应的输出 y 可能代价高昂。因此，通常需要仅通过少量 y观测值来执行 (1)。为此，我们针对 f采用了结合高斯过程建模 [28] 的主动学习方法。

建模与算法 f 的零均值高斯过程模型 $\mathcal{GP}(0, k(x, x'))$ 由核函数 $k(x, x')$ 表征。核函数将关于 f 平滑性的先验知识纳入模型，并评估输入之间的相似性。在此

研究中，我们使用了径向基函数（RBF）核，其定义如下：

$$k(x, x') = v \exp\left(-\frac{\|x - x'\|^2}{2\ell^2}\right), \tag{2}$$

其中，v 和 ℓ 是需要适当调节的超参数。假设当前处于第 t 步，且已观测到输入和输出对的数据 $\mathcal{D}_t = \{(x_i, y_i)\}_{i=1}^{N_t}$。那么，在观测到数据 $\mathcal{D}_N$的条件下，新输入 x 对应的目标值 $f(x)$的预测分布为正态分布 $\mathcal{N}(\mu_t(x)\ \sigma_t^2(x))$，其预测均值为$\mu_t(x)$，预测方差为 $\sigma_t^2(x)$，定义如下：

$$\mu_t(x) = \boldsymbol{k}(x)^\top (\boldsymbol{K} + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{y}, \tag{3}$$

$$\sigma_t^2(x) = k(x, x) + \boldsymbol{k}(x)^\top (\boldsymbol{K} + \sigma^2 \boldsymbol{I})^{-1} \boldsymbol{k}(x), \tag{4}$$

其中，$\boldsymbol{k}(x) = (k(x, x_1) ..., k(x, x_{N_t}))$，$\boldsymbol{K}$是一个大小为 $N_t \times N_t$的矩阵，其第(i, j)个元素定义为 $k(x_i, x_j)$，i、$j = 1$、...、N_t，且 $\boldsymbol{I}$是单位矩阵。基于这一建模，我们的LSE主动学习算法按以下步骤执行：(i) 定义置信区域 $Q_t(x) = [\mu_t(x) \pm 1.96\sigma_t(x)]$。(ii) 如果 $\min Q_t(x) + \varepsilon = \mu_t(\boldsymbol{x}) - 1.96\sigma_t(\boldsymbol{x}) + \varepsilon \geq \theta$，则将候选点 $x \in \mathcal{X}$ 分类为 $x \in \mathcal{U}_\theta$；如果 $\max Q_t(x) - \varepsilon = \mu_t(\boldsymbol{x}) + 1.96\sigma_t(\boldsymbol{x}) - \varepsilon < \theta$，则将其分类为 $x \in \mathcal{L}_\theta$。(iii) 通过最大化采集函数（称为跨越准则）[8] 来选择下一个待评估点，其定义如下：

$$x_{N_t+1} = \arg\max_{x \in \mathcal{X}} [\min\{\max Q_t(x) - \theta, \theta - \min Q_t(x))\}] = \underbrace{\arg\max_{x \in \mathcal{X}} 1.96\sigma_t(x) - |\mu_t(x) - \theta|}_{\text{跨越准则}} \tag{5}$$

(iv) 使用数据 $\mathcal{D}_{t+1} = \mathcal{D}_t \cup \{(x_{N_t+1}, y_{N_t+1})\}$更新高斯过程模型。算法的详细信息见补充信息1。图3展示了使用合成数据进行 LSE 主动学习的示例。

基于主动学习的 LSE 与 Diff-GP 模型 考虑这样一种情况：除当前实验的数据外，还获得了来自相似实验 f' 的数据集 $\mathcal{D}' = \{(x'_j, y'_j)\}_{j=1}^{M}$。Diff-GP 模型 [32] 是一种对 $\mathcal{D}'$ 进行变换的方法，使其可被视为当前实验的数据。变换后的 D' 与当前实验的数据合并，并用于建模。简而言之，在 Diff-GP 模型中，利用均值函数$\mu_t(x)$ f 的高斯过程对当前实验 f 与相似实验 f' 之间的差异 $\Delta f = f - f'$ 建模；随后，将变换后的数据集 $\hat{}$ $\mathcal{D} = \{x'_j, \hat{} y_j = y'_j + \mu_t(x'_j)\}_{j=1}^{M}$ 视为当前实验的$\hat{}$数据。然后，使用 $\mathcal{D}_t \cup \mathcal{D}$ 构建高斯过程模型，作为

一个数据集，并按照上一节所述，使用主动学习自适应地执行 LSE。算法的详细信息见补充信息 2。

位置-尺度偏移下基于主动学习的 LSE 与 Diff-GP 模型 Diff-GP 模型是一种隐含假设 $\mathcal{D}_t$ 和 $\mathcal{D}'$ 相似，并使用高斯过程对 f 和 f' 之间的差异进行建模的方法。在此，如果实验环境控制不当，输出的位置和尺度可能会发生显著波动（即上述假设不成立），因此 Diff-GP 模型无法直接应用。然而，即使这一假设不成立，仍存在一些情况，可以将 $\mathcal{D}'$ 转换为与 $\mathcal{D}_t$ 相似的形式，从而应用 Diff-GP 模型。在本研究中，我们考虑这样一种情况： f' 的 位置-尺度偏移 [38] 版本可以表示为

$$f''(x) = \gamma f'(x) + \eta, \tag{6}$$

这样，我们就可以将 Diff-GP 模型应用于 f 和 f''。其中， γ 和 η 分别是尺度偏移参数和位置偏移参数，由观测数据估计得到。如果估计出 γ 和 η ˜，则根据公式（6）对 $\mathcal{D}'$ 进行变换所获得的数据集 $\mathcal{D}$ 会与$\mathcal{D}_t$合并，然后应用上一节所述的 Diff-GP 模型。算法的详细信息见补充信息 3。

非自适应基线方法 给定一组预设测量点后，通常会如图1b所示，从材料表面的边缘开始依次进行测量。然而，采用这种方法时，有些区域要等到所有测量完成后才会被检查，而且作为基线方法，其表现较差。因此，我们考虑了一种基于二维材料表面递归划分的基线方法。对于该基线方法，如补充图1所示，以已测量点为边界递归生成的矩形，其中心被选作新的测量点。

报告摘要 有关实验设计的更多信息，请参阅本文附带的 Nature Research 报告摘要。

数据可用性 如有合理请求，可向通讯作者获取支持本研究发现的数据。

参考文献

[1] N Balke, S Jesse, AN Morozovska, E Eliseev, DW Chung, Y Kim, L Adamczyk, RE Garcia, N Dudney, and SV Kalinin. 锂离子电池正极中离子扩散的纳米尺度映射。《自然·纳米技术》，5(10):749–754，2010年。[2] Dale E Newbury and Nicholas WM Ritchie. 使用扫描电子显微镜-能量色散X射线谱分析（SEM-EDS）对微观结构进行元素映射：硅漂移探测器（SDD）带来的非凡进展。《分析原子光谱学杂志》，28(7):973–988，2013年。[3] Reiner Salzer and Heinz W Siesler. 红外与拉曼光谱成像。John Wiley & Sons，2014年。[4] David A Shapiro, Young-Sang Yu, Tolek Tyliszczak, Jordi Cabana, Rich Celestre, Weilun Chao, Konstantin Kaznatcheev, AL Kilcoyne, Filipe Maia, Stefano March- esini, et al. 通过软X射线显微术实现纳米级分辨率的化学成分映射。《自然·光子学》，8(10):765–769，2014年。[5] Genxiang Zhong, Qinghua Yu, Xinming Huang, and Lijun Liu. 影响籽晶辅助铸锭底部低少数载流子寿命区形成的因素。Journal of crystal growth，402:65–70，2014年。[6] Dieter K Schroder. 半导体材料与器件表征。John Wiley & Sons，2015年。[7]Ville V¨ah¨anissi, Hannu S Laine, Zhengjun Liu, Marko Yli-Koski, Antti Haarahiltunen, and Hele Savin. p型高性能多晶硅中红色区域的完全恢复。Solar Energy Materials and Solar Cells，173:120–127，2017年。[8]B. Bryan, R. C. Nichol, C. R. Genovese, J. Schneider, J. M. Christopher, and L. Wasserman. 用于识别函数阈值边界的主动学习。载于《神经信息处理系统进展》，第163–170页，2006年。[9] Alkis Gotovos, Nathalie Casati, Gregory Hitz, and Andreas Krause. 用于水平集估计的主动学习。载于《第二十三届国际人工智能联合会议》，2013年。[10] Andrea Zanette, Junzi Zhang, and Mykel J Kochenderfer. 使用高斯过程进行稳健的上水平集估计。载于《机器学习与数据库知识发现联合欧洲会议》，第276–291页。Springer，2018年。

[11] Shubhanshu Shekhar 和 Tara Javidi。多尺度高斯过程水平集估计。arXiv 预印本 arXiv:1902.09682，2019 年。[12] Shogo Iwazaki 、 Yu Inatsu 和 Ichiro Takeuchi。输入不确定性下寻找可靠水平集的贝叶斯实验设计。IEEE Access，8：203982–203993，2020年。[13] Yu Inatsu、Masayuki Karasuyama、Keiichi Inoue 和 Ichiro Takeuchi。针对输入不确定性下水平集估计的主动学习及其扩展。Neural Computation，32(12)：2486–2531，2020年。[14]Burr Settles。主动学习文献综述。技术报告，威斯康星大学麦迪逊分校计算机科学系，2009年。[15] Burr Settles。从理论到查询：实践中的主动学习。载于：与 AISTATS 2010 联合举办的主动学习与实验设计研讨会，第1–18页，2011年。[16] Jasper Snoek、Hugo Larochelle 和 Ryan P Adams。机器学习算法的实用贝叶斯优化。载于：神经信息处理系统进展，第2951–2959页，2012年。[17] Bobak Shahriari、Kevin Swersky、Ziyu Wang、Ryan P Adams 和 Nando De Freitas。让人类退出循环：贝叶斯优化综述。IEEE 论文集，104(1)：148–175，2015年。[18] Yu Inatsu、Daisuke Sugita、Kazuaki Toyoura 和 Ichiro Takeuchi。基于高斯过程导数枚举局部极小值的主动学习。Neural Computation，32(10)：2032–2068，2020年。[19] Kathryn Chaloner 和 Isabella Verdinelli。贝叶斯实验设计：综述。Statistical Science，第273–304页，1995年。[20] Jerry P Queen、Gerry P Quinn 和 Michael J Keough。生物学家的实验设计与数据分析。剑桥大学出版社，2002年。[21] Friedrich Pukelsheim。实验的最优设计。SIAM，2006年。[22] Roger E Kirk。实验设计。《心理学手册》第二版，第2卷，2012年。

[23] Sinno Jialin Pan 和 Qiang Yang。迁移学习综述。IEEE Transactions on Knowledge and Data Engineering，22(10)：1345–1359，2009年。[24] Kentaro Kutsukake、Momoko Deura、Yutaka Ohno 和 Ichiro Yonenaga。硅锭表征：类单晶与高性能多晶的比较。Japanese Journal of Applied Physics，54(8S1)：08KD10，2015年。[25] Huali Zhang、Da You、Chunlai Huang、Yihua Wu、Yan Xu 和 Peng Wu。通过石英籽晶法生长兼具更高品质和产量的多晶硅锭。Journal of Crystal Growth，435：91–97，2016年。[26] Dongli Hu、Shuai Yuan、Xuegong Yu、Liang He、Yunfei Xu、Xueri Zhang 和 Deren Yang。高性能多晶硅的晶界工程：控制硅锭边缘的铁污染。Solar Energy Materials and Solar Cells，171：131–135，2017年。[27] Zhong Li、Jia-Dan Li、Lin Zhuang 和 Rui-Jiang Hong。采用籽晶辅助定向凝固生产的高性能多晶硅锭外围边缘区域的少数载流子寿命评估。Frontiers of Physics，12(5)：128103，2017年。[28]Christopher KI Williams 和 Carl Edward Rasmussen。用于机器学习的高斯过程。麻省理工学院出版社，马萨诸塞州剑桥，2006年。[29] Richard O Duda、Peter E Hart 和 David G Stork。模式分类。John Wiley & Sons，2012年。[30] Masashi Sugiyama、Taiji Suzuki 和 Takafumi Kanamori。机器学习中的密度比估计。剑桥大学出版社，2012年。[31] Nicolas Courty、R´emi Flamary、Devis Tuia 和 Alain Rakotomamonjy。用于领域自适应的最优传输。IEEE Transactions on Pattern Analysis and Machine Intelligence，39(9)：1853–1865，2017年。[32]A. Shilton、S. Gupta、S. Rana 和 S. Venkatesh。贝叶斯优化中迁移学习的遗憾界。载于：AISTATS 2017：Machine Learning Research：第20届人工智能与统计国际会议论文集，第1–9页，2017年。[33] Marc G Genton。机器学习核函数类别：统计学视角。Journal of Machine Learning Research，2(12月)：299–312，2001年。

[34] Randy Jalem, Kenta Kanamori, Ichiro Takeuchi, Masanobu Nakayama, Hisatsugu Yamasaki, and Toshiya Saito. 贝叶斯驱动的第一性原理计算，用于加速探索适用于可充电电池的快离子导体。Scientific Reports, 8(1):1–10, 2018年。

[35] Shali Jiang, Gustavo Malkomes, Geoff Converse, Alyssa Shofner, Benjamin Moseley, and Roman Garnett. 高效的非短视主动搜索。载于《第34届国际机器学习会议论文集》第70卷，第1714–1723页。JMLR.org，2017年。

[36] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. 强化学习：综述。《人工智能研究杂志》，4:237–285，1996。

[37] Richard S Sutton, Andrew G Barto, et al. 强化学习导论，第135卷。麻省理工学院出版社，剑桥，1998。

[38] I. Takeuchi、K. Nomura 和 T. Kanamori. 使用核分位数回归的分段线性解路径进行非参数条件密度估计. Neural Computation, 21(2):533–559, 2009年。

致谢

本研究获得了日本学術振興会（JSPS）科学研究科研费（KAKENHI，项目编号：20K19871 和 19H04071，资助 K.M.；20H00601，资助 I.T.）、日本同步辐射研究所（JASRI）（提案编号：2018A3585、2018B3587、2019A3587、2019B3586 和 2020A3586，资助 K.K.）、日本科学技术振兴机构（JST）CREST（项目编号：JP-MJCR21D3）、JST Moonshot研发计划（项目编号：JPMJMS2033-05）、JST AIP加速研究（项目编号：JPMJCR21U2）以及新能源与产业技术综合开发机构（NEDO，项目编号：JPNP18002 和 JPNP20006，资助 I.T.）的支持；K.K.、K.M.、T.U. 和 I.T. 获得了理化学研究所先进智能项目中心的支持。

作者贡献

K.K.、K.M.、T.U. 和 I.T. 参与了研究设计。K.K. 收集、管理并分析了太阳能电池铸锭数据。H.S.、K.M.、K.K. 和 I.T. 构建了机器学习方法（基于AL的LSE和基于 ATL 的 LSE 方法）。H.S. 和 S.K. 实施了这些方法并进行了数值实验。H.S.、K.K、K.M. 和 I.T. 撰写了稿件。所有作者均讨论并评论了稿件。

利益冲突

作者声明不存在利益冲突。

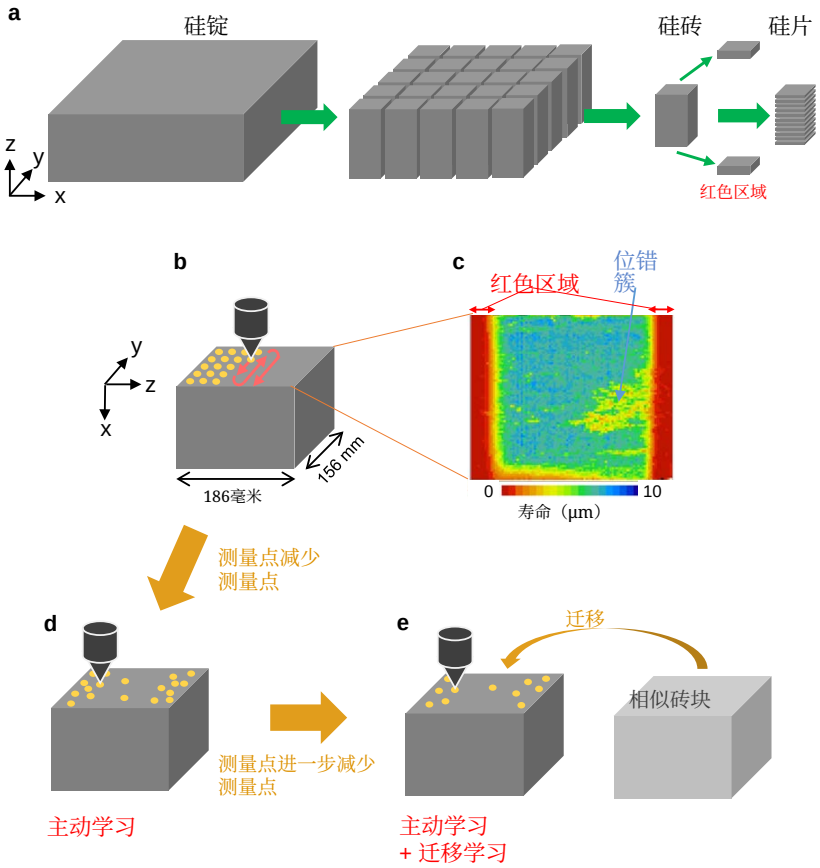


图1：太阳能电池硅锭红区识别问题示意图。**a** 由硅锭制造多晶硅片的工艺。切片工艺开始前，先从硅砖中去除红色区域。**b** 在 y-z 砖块侧面的网格点上逐点测量硅锭截面载流子寿命的传统空间映射。**c** 载流子寿命的典型空间分布。**d**通过基于主动学习（AL）的自适应空间映射减少测量点。**e**利用先前生产的相似砖块中的信息，通过迁移学习（TL）进一步减少测量点。

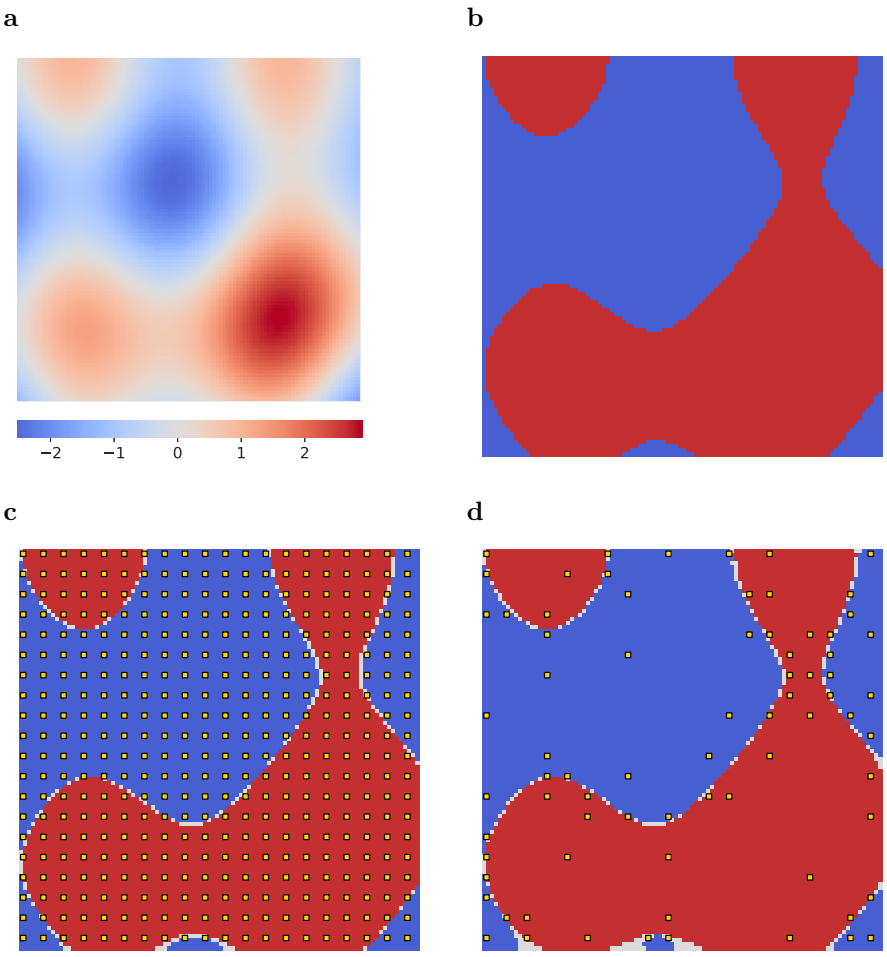


图2：二维函数 f 的水平集估计（LSE）问题示意图；颜色表示定义在二维输入空间 $\mathcal{X}$ 上的函数值。**a** 材料表面物理性质真实空间分布示例。**b** 阈值 θ 对应的超水平集和次水平集分别定义为 $\mathcal{U}_\theta = \{\boldsymbol{x} \in \mathcal{X} \mid f(\boldsymbol{x}) > \theta\}$和 $\mathcal{L}_\theta = \{\boldsymbol{x} \in \mathcal{X} \mid f(\boldsymbol{x}) \leq \theta\}$。对于图1中的问题，红区表示职业生涯时长给定阈值所对应的次水平集。**c**传统水平集估计采用穷举测量，对所有预先设定的网格点进行测量。一旦评估了所有预设网格点处的函数值，就能准确估计超水平集和次水平集；但这种方法通常效率不高。**d**基于主动学习（AL）的水平集估计在选定的网格点进行自适应测量。即使仅评估缺陷区域边界附近选定网格点处的函数，也能准确估计超水平集 $\mathcal{U}_\theta$ 和 $\mathcal{L}_\theta$。



图3：使用基于主动学习的水平集估计方法，对示例一维函数 f 进行缺陷区域识别的示意图。从左上到右下的各图展示了截至当前迭代获得的训练数据点（方块），以及每次迭代的贝叶斯回归模型。拟合贝叶斯回归（高斯过程）模型后，可得到每个输入位置 $x \in \mathcal{X}$ 处的均值$\mu(x)$（深蓝色曲线）和置信区间 $\mu(x) \pm 1.96\sigma(x)$（浅蓝色区域）。该水平集估计问题的目标是识别函数值大于 θ （红色虚线）的输入位置子集；每幅图中的红色阴影区域表示估计的缺陷区域。此示例表明，输入点更有可能在函数值接近阈值的位置被选取。

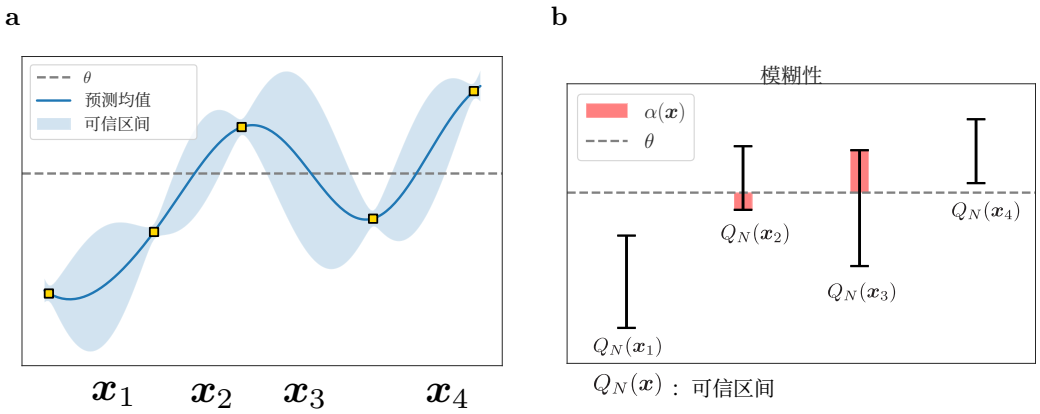
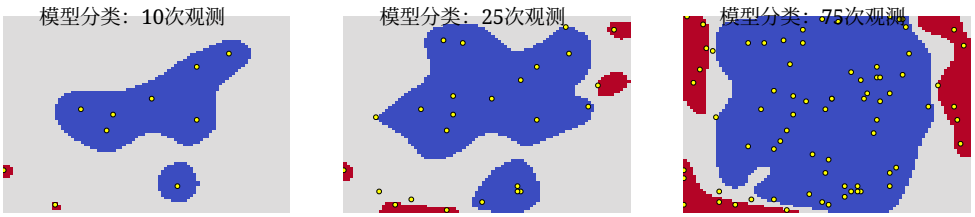


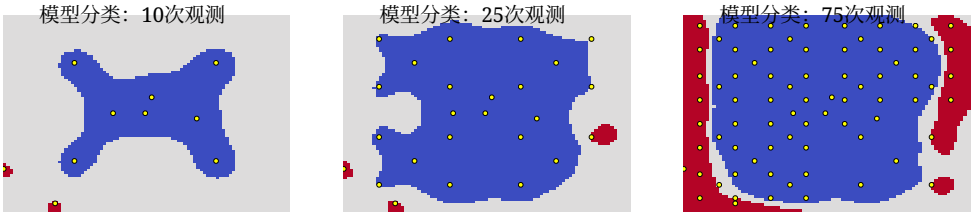
图4：基于AL的LSE方法针对示例性一维函数 f 的自适应选择准则示意图。**a** 某次迭代时的贝叶斯回归模型。**b** 对于四个输入候选点 x_1，x_2，x_3，x_4，该方法选择 x_3，因为其违反度$\alpha(x_3)$（浅红色区域）在四个候选点中最大。每个输入候选点处的违反度定义为 $\alpha(x) = \min\{\max\{0, \theta - \ell(x)\}, \max\{0, u(x) - \theta\}\}$ 其中 $\ell(x)$ 和 $u(x)$ 分别是可信区间 $Q_N(x)$ 的下限和上限。

a



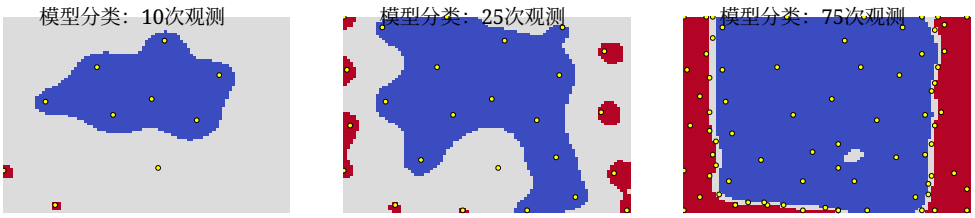
随机方法

b



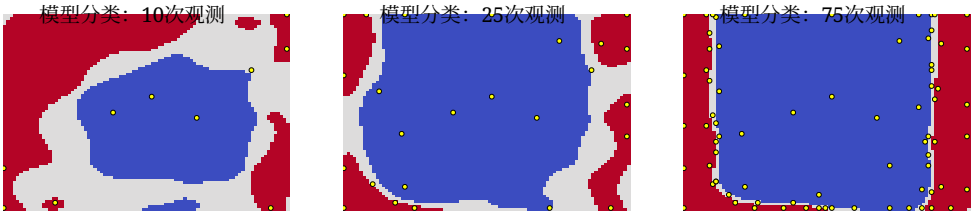
非自适应方法

c



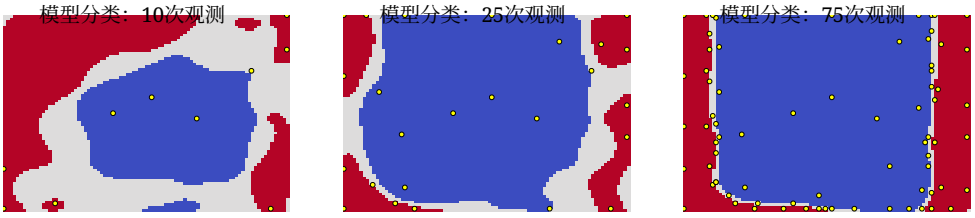
基于主动学习的水平集估计方法

d



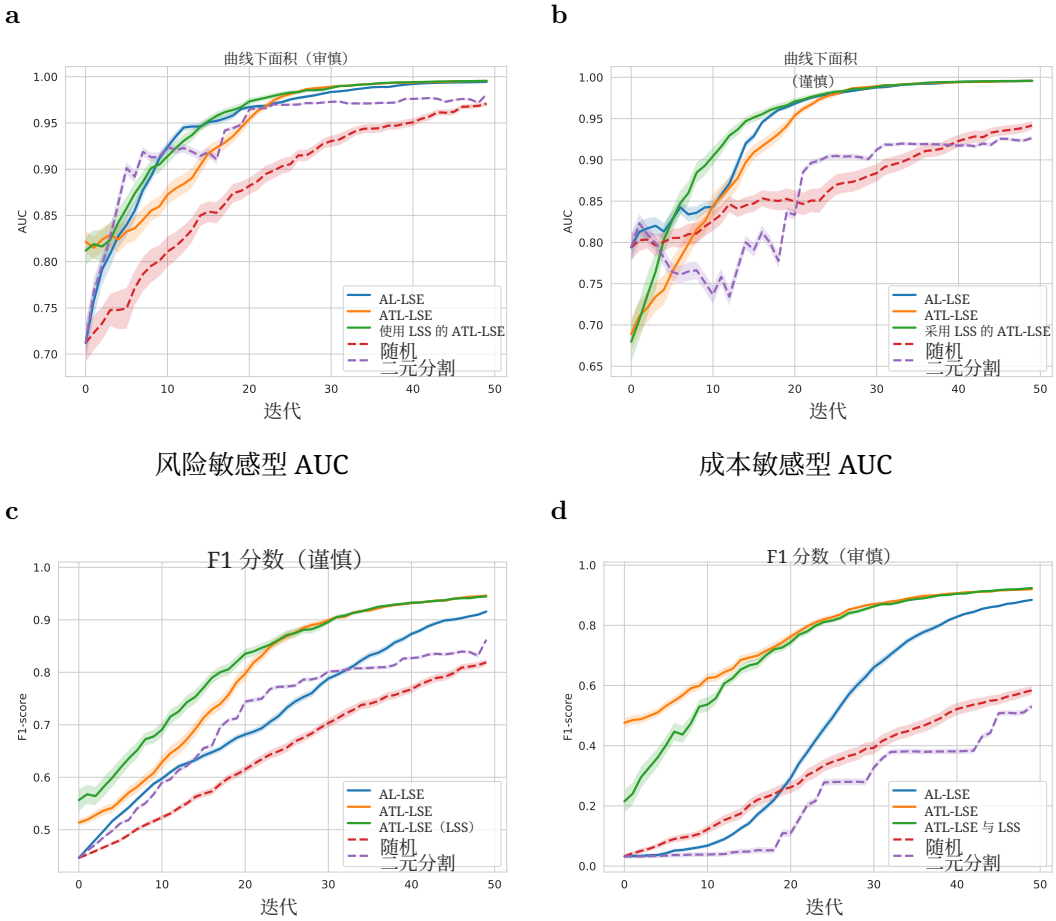
基于 ATL 的 LSE 方法

e



基于 LSS-ATL 的 LSE 方法

图 5：比较两种基线方法（随机和非自适应方法）与三种所提出的方法（基于 AL-、ATL- 和 LSS-ATL 的 LSE），用于识别太阳能电池硅锭中的红区。每种方法中的三幅图分别表示进行 10、25 和 75 次测量后测得的测量点，以及估计的上水平集（蓝色）、次水平集（红色）和不确定区域（灰色）。



风险敏感型 AUC　　成本敏感型 AUC

风险敏感型 F1 分数　　成本敏感型 F1 分数

图 6：比较两种基线方法（随机和非自适应二元分割）与三种所提出的方法（基于 AL-、ATL- 和 LSS-ATL 的 LSE）在风险敏感和成本敏感情境下的曲线下面积（AUC）和 F1 分数。

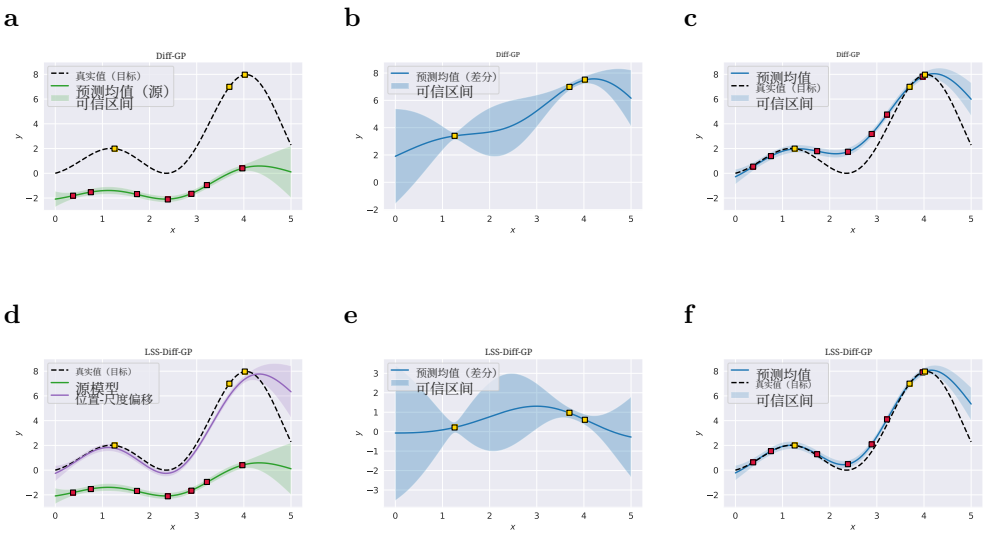
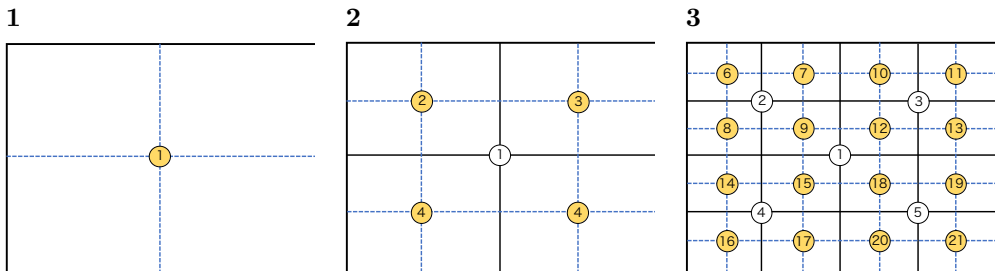


图 7：基于 ATL 的 LSE 方法的基本概念（**a–c**）和位置-尺度偏移示例（**d–f**）。**a** 源域（绿色）的观测值和单独的高斯过程模型。**b**目标函数与源函数之差的高斯过程模型。**c** 目标域的高斯过程模型，以及从源域迁移的知识。**d** 源域（绿色）的观测值和单独的高斯过程模型，以及经过位置-尺度偏移变换后的函数（紫色）。**e** 位置-尺度偏移目标函数与源函数之差的高斯过程模型。**f** 位置-尺度偏移下目标域的高斯过程模型，以及从源域迁移的知识。



补充图1 非自适应方法示意图。

算法 1 基于主动学习的水平集估计（基于AL的LSE）

输入： 观测数据 $\mathcal{D}_t$、高斯过程先验、一组候选点 $\mathcal{X}$、最大迭代次数 T

输出： $\mathcal{U}_\theta$， $\mathcal{L}_\theta$， $\mathcal{C}_\theta$

初始化： $\mathcal{U}_\theta$， $\mathcal{L}_\theta \leftarrow \emptyset$， $\mathcal{C}_\theta \leftarrow \mathcal{X}$， $t \leftarrow 0$

对于 $t = 1$、2、.....、T**执行**

步骤 1 运行高斯过程回归，得到式（3）中定义的预测均值 $\mu_t(\boldsymbol{x})$ 和预测方差 $\sigma_t(\boldsymbol{x})$。

步骤 2 将置信区域定义为 $Q_t(\boldsymbol{x}) = [\mu_t(\boldsymbol{x}) - 1.96\sigma_t(\boldsymbol{x})\ \mu_t(\boldsymbol{x}) + 1.96\sigma_t(\boldsymbol{x})]$**步骤 3** 按如下方式对所有候选点 $\boldsymbol{x} \in \mathcal{X}$ 进行分类:

$$\begin{cases} \mathcal{U}_\theta \leftarrow \mathcal{U}_\theta \cup \{\boldsymbol{x}\} & \text{if } \mu_t(\boldsymbol{x}) - 1.96\sigma_t(\boldsymbol{x}) + \varepsilon \geq \theta, \\ \mathcal{L}_\theta \leftarrow \mathcal{L}_\theta \cup \{\boldsymbol{x}\} & \text{if } \mu_t(\boldsymbol{x}) + 1.96\sigma_t(\boldsymbol{x}) - \varepsilon < \theta, \end{cases}$$

并将不确定集更新为 $\mathcal{C}_\theta \leftarrow \mathcal{X} \backslash \mathcal{U}_\theta \cup \mathcal{L}_\theta$。

步骤 4 按如下方式，通过最大化跨越采集函数选择下一个评估点:

$$\boldsymbol{x}_{i(t)} = \arg\max_{\boldsymbol{x} \in \mathcal{X}} 1.96\sigma_t(\boldsymbol{x}) - |\mu_t(\boldsymbol{x}) - \theta|$$

步骤 5 将数据集更新为 $\mathcal{D}_t \leftarrow \mathcal{D}_t \cup \{(\boldsymbol{x}_{i(t)}, y_{i(t)})\}$**对所有项执行完毕**

补充信息1：LSE 主动学习（AL）算法详解

本节总结提出的三种算法：基于AL的LSE、基于ATL的LSE和LSS-基于ATL的LSE。令 $\mathcal{D}_t = \{(x_i,\ y_i)\}_{i \in S_t}$ 表示步骤 t时的已观测数据集，其中 $S_t \subset \{1, ..., N\}$。我们的目标是将所有候选点 $\boldsymbol{x} \in \mathcal{X}$划分为上水平集 $\mathcal{U}_\theta$、次水平集 $\mathcal{L}_\theta$或未确定集 $\mathcal{C}_\theta$。请注意，$\mathcal{C}_\theta$ 对于恰当考虑问题的不确定性是必要的。算法 1 给出了基于AL的LSE算法的伪代码。首先，我们对步骤 1 中已观测的数据进行高斯过程回归，以获得预测均值函数和方差函数。在步骤 2 和步骤 3 中，我们使用步骤 1 中得到的预测均值函数和方差函数所构成的可信区间，对候选点进行分类。然后，在步骤 4 中最大化跨越采集函数，并确定下一个观测点。最后，将观测数据添加到数据集中，并返回步骤 1。该过程不断循环，直到所有候选点均被划分为 $\mathcal{U}_\theta$ 或 $\mathcal{L}_\theta$，或者达到最大迭代次数。

算法 2 通过 Diff-GP 模型进行数据变换

输入： 观测源数据 $\mathcal{D}'$、高斯过程先验、一组候选点 $\mathcal{X}$、最大迭代次数 T

输出： 变换后的源数据 $\hat{\mathcal{D}} = \{(\boldsymbol{x}'_j,\ \hat{}y_j)\}_{j=1}^{M}$

步骤 1 使用 $\mathcal{D}'$ 执行高斯过程回归，以获得预测均值函数 $\mu'(\boldsymbol{x})$ 和预测方差函数 $\sigma'^2(\boldsymbol{x})$。

步骤 2 按如下方式计算 Diff-GP 模型的预测均值和方差：

$$\hat{\mu}(x) = \boldsymbol{k}(\boldsymbol{x})^\top \left(\boldsymbol{K} + (\sigma^2 + \sigma'^2(\boldsymbol{x}))\boldsymbol{I}\right)^{-1} (\boldsymbol{y} - \boldsymbol{\mu}'),$$
$$\hat{\sigma}^2(x) = k(\boldsymbol{x}, \boldsymbol{x}) + \boldsymbol{k}(\boldsymbol{x})^\top \left(\boldsymbol{K} + (\sigma^2 + \sigma'^2(\boldsymbol{x}))\boldsymbol{I}\right)^{-1} \boldsymbol{k}(\boldsymbol{x}),$$

其中，$\boldsymbol{k}(\boldsymbol{x})$ 和 $\boldsymbol{K}$ 分别是根据目标数据定义的核向量和核矩阵，σ^2 是目标数据的观测误差，而 $\boldsymbol{\mu}' = (\mu'(\boldsymbol{x}_1) ..., \mu'(\boldsymbol{x}_n))$。**步骤 3** 计算变换后的输出，即 $\hat{}y_j = y'_j + \hat{}\mu(\boldsymbol{x}'_j)$，$j = 1, ..., M$。

补充信息 2: : LSE 的主动迁移学习（ATL）算法详情

对于基于 ATL 的 LSE 方法，算法 2 所示的数据变换步骤插入在算法 1 的步骤 1 之前。设 $\mathcal{D}' = \{(\boldsymbol{x}'_j, y'_j)\}_{j=1}^{M}$ 为源域的已观测数据集。Diff-GP 模型是一种算法，可将源域数据变换为可视为目标域数据的数据。算法 2 的输出（即变换后的数据集 $\mathcal{D}$）与目标数据集 $\mathcal{D}_t$合并，并用作算法 1 的数据集。

补充信息 3：用于 LSE 的位置尺度偏移主动迁移学习（ LSS-ATL）算法详情

对于基于 LSS-ATL 的 LSE 方法，在算法 1 的步骤 1 之前插入算法 3 所示的数据变换步骤。算法 3 结合了 Diff-GP 模型和位置尺度偏移变换。在本研究中，我们假设位置-尺度偏移是一个包含两个参数的仿射变换： γ 和 η。我们通过从观测数据中估计 γ 和 η 来变换源数据。

算法 3 通过 LSS-Diff-GP 模型进行数据变换

输入： 观测源数据 $\mathcal{D}'$、高斯过程先验、一组候选点 $\mathcal{X}$、最大迭代次数 T

输出： 变换后的源数据 $\hat{\mathcal{D}} = \{(\boldsymbol{x}'_j, \hat{y}_j)\}_{j=1}^{M}$

步骤 1 使用 $\mathcal{D}'$ 运行高斯过程回归，得到预测均值函数 $\mu'(\boldsymbol{x})$ 和预测方差函数 $\sigma'^2(\boldsymbol{x})$。

步骤 2 求解关于 γ 和 η的以下最小二乘问题：

$$\gamma^*, \eta^* = \arg\min_{\gamma,\eta} \sum_{i=1}^{n} \left(y_i - (\gamma\mu'(\boldsymbol{x}_i) + \eta)\right)^2.$$

步骤 3 按如下方式计算 Diff-GP 模型的预测均值和方差：

$$\hat{\mu}(x) = \boldsymbol{k}(\boldsymbol{x})^\top \left(\boldsymbol{K} + (\sigma^2 + \sigma'^2(\boldsymbol{x}))\boldsymbol{I}\right)^{-1} (\boldsymbol{y} - (\gamma^*\boldsymbol{\mu}' + \eta^*)),$$
$$\hat{\sigma}^2(x) = k(\boldsymbol{x}, \boldsymbol{x}) + \boldsymbol{k}(\boldsymbol{x})^\top \left(\boldsymbol{K} + (\sigma^2 + \sigma'^2(\boldsymbol{x}))\boldsymbol{I}\right)^{-1} \boldsymbol{k}(\boldsymbol{x}),$$

其中，$\boldsymbol{k}(\boldsymbol{x})$ 和 $\boldsymbol{K}$ 分别是根据目标数据定义的核向量和核矩阵，σ^2 是目标数据的观测误差，而 $\gamma^*\boldsymbol{\mu}' + \eta^* = (\gamma^*\mu'(\boldsymbol{x}_1) + \eta^*, ..., \gamma^*\mu'(\boldsymbol{x}_n) + \eta^*)$。

步骤 4 计算变换后的输出：$\hat{y}_j = y'_j + \hat{\mu}(\boldsymbol{x}'_j)$，$j = 1, ..., M$。