

文章编号: 1000-7032(2024)03-0399-08

机器学习辅助钙钛矿薄膜制备工艺优化及特征重要性评估

弓箭¹, 陈谦¹, 李阳^{1*}, 马梦恩²,
马玉姣², 吴绍航², 刘冲^{2*}, 麦耀华²

(1. 五邑大学 智能制造学部, 广东 江门 529020;

2. 暨南大学 物理与光电工程学院, 广东 广州 510632)

摘要: 钙钛矿太阳能电池仅用十年左右的时间将效率提升至认证的26.1%, 非常接近晶硅太阳能电池26.81%的认证效率, 展现出巨大的产业化潜力。当前, 钙钛矿太阳能电池器件效率还在提升, 然而在器件制备过程中, 钙钛矿太阳能电池的性能受到许多不可分割的因素影响, 传统方法往往采用试错的方式来优化钙钛矿太阳能电池的制备工艺, 花费了大量的时间。贝叶斯优化是一种全局优化算法, 在解决人工智能的黑盒问题方面取得了很大的成功。本文利用贝叶斯优化算法对钙钛矿层涉及到的碘化铅(PbI₂)过量百分比、退火温度、退火时间、真空萃取时间四个工艺参数进行优化选择, 显著降低了研发成本, 缩短了研发时间。通过五轮实验迭代, 累计34组工艺条件, 制备出了器件效率为23.56%的反型钙钛矿太阳能电池。

关键词: 钙钛矿太阳能电池; 机器学习; 工艺优化; 高效率

中图分类号: O482.31; TM914.4 文献标识码: A DOI: 10.37188/CJL.20230309

Machine Learning Assisted Optimization of Perovskite Thin Film Fabrication Process and Assessment of Feature Importance

GONG Jian¹, CHEN Qian¹, LI Yang^{1*}, MA Meng'en², MA Yujiao²,
WU Shaohang², LIU Chong^{2*}, MAI Yaohua²

(1. Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China;

2. College of Physics and Optoelectronic Engineering, Jinan University, Guangzhou 510632, China)

* Corresponding Authors, E-mail: insidesun51@163.com; chongliu@jnu.edu.cn

Abstract: The efficiency of perovskite solar cells has been improved to 26.1% in just ten years, which is very close to the certification efficiency of crystalline silicon solar cells (26.81%). This demonstrates the significant potential for industrialization. Currently, efforts are still being made to further enhance the efficiency of perovskite solar cells. However, various inseparable factors affect the performance of perovskite solar cells during the device preparation process. Traditional methods often rely on trial and error to optimize the preparation process, resulting in time-consuming procedures. Bayesian optimization, a global optimization algorithm, has achieved remarkable success in addressing artificial intelligence's black box problem. In this work, the Bayesian optimization is employed to optimize four key process parameters involved in the perovskite layer: excess percentage of lead iodide (PbI₂), annealing temperature, annealing time, and vacuum extraction time. The costs of research and development have been significantly reduced, as well as the required time for such activities has also been shortened. The improvement was achieved through five rounds of experimental iterations and 34 sets of process conditions, ultimately resulting in the preparation of an inverse perovskite solar cell with a device efficiency rating of 23.56%.

收稿日期: 2023-12-02; 修订日期: 2023-12-19

基金项目: 广东省重点领域研发计划项目(2019B010132004); 国家自然科学基金(62104082, 62005099); 广东省基础与应用基础研究基金(2022A1515010746, 2022A1515011228); 广州市科技计划(202201010458, 202201010542)

Supported by The Key Realm R&D Program of Guangdong Province(2019B010132004); National Natural Science Foundation of China(62104082, 62005099); Guangdong Basic and Applied Basic Research Foundation(2022A1515010746, 2022A1515011228); Science and Technology Program of Guangzhou(202201010458, 202201010542)

Key words: perovskite solar cells; machine learning; process optimization; high efficiency

1 引 言

钙钛矿太阳能电池因其光吸收系数高、带隙可调、载流子扩散长度长、电荷载流子迁移率高、制造成本低等特点引起了人们的极大关注。自2009年发明以来^[1],能量转换效率已经从3.8%提高到认证的26.1%,被认为是最有前途的下一代光伏技术。

钙钛矿光吸收层作为钙钛矿太阳能电池的重要组成部分,是器件性能的决定性因素。因此,制备形貌可控、高表面覆盖率和少针孔的高质量钙钛矿薄膜就成为制备过程中的一个重要环节。影响钙钛矿薄膜成膜质量的因素还包括溶液浓度、前驱体组成、添加剂、成膜温度和时间等,这其中不同的工艺条件对工艺窗口的影响不尽相同^[2]。因此,需要采取大量的时间和精力,采用控制变量的方法进行逐项摸索。然而更为关键的是,多种工艺条件往往不是独立影响,而是各个参数间相互影响的,这种情况下,传统手段采用的控制单一变量的方法来筛选最佳实验条件往往不能达到最佳效果。大数据和人工智能被称为“第四次工业革命”,机器学习是人工智能的重要分支之一,是一种让计算机学习如何从已知数据中执行特定任务的技术。由于数据的快速积累和计算机硬件的进步,已成为近年来的研究热点,在化学、材料以及制造业领域的应用越来越多^[3-5]。对于钙钛矿太阳能电池制造工艺来说,由于广泛的加工参数和不同性能之间的权衡,很难确定光伏材料的期望性能与可控的制造条件之间的确切关系。机器学习可以根据需求创建模型,这些模型可以使用现有数据有效地训练。对于太阳能电池制备工艺优化问题,函数往往呈现多维,表达式未知,可能存在很多局部最优解等复杂情况。贝叶斯优化方法则采用贝叶斯思想,通过不断探索各种参数组合的结果,根据已有信息计算期望值,并选择期望值最大的组合作为最佳策略,从而在尽可能少的实验次数下达到最优解^[6]。

机器学习在太阳能电池设计和制造中的应用起步缓慢,但最近获得了巨大的发展势头,不断加速新的钙钛矿太阳能电池的探索和发展。Li等回顾了机器学习在太阳能电池制造中的应用,并根据其目标将这些工作分为四类:材料性能的预测、

器件结构的优化、制造工艺的优化以及测量数据的重建^[7]。严文生教授团队利用机器学习和光管理实现了效率为23.6%的钙钛矿太阳能电池^[8]。Zhang等建立高斯过程回归(Gaussian process regression, GPR)模型来阐明立方钙钛矿 A_2XY_6 化合物的离子半径和晶格常数之间的统计关系^[9]。Li和Tao等构建随机森林分类(Random forest classification, RFC)模型来区分钙钛矿对 ABX_3 和 $A_2B'B''X_6$ 化合物的可成形性^[10]。

Sun和Buonassisi团队在2个月内制造并表征了75种独特的钙钛矿型组合物,并利用完全连接的深度神经网络将实测的化合物分类为0D、2D和3D结构,准确率高达90%^[11]。Vakharia等利用弹性网络回归和保序回归模型对金属卤化物钙钛矿带隙进行预测^[12]。中国香港大学Wang教授通过将机器学习方法与密度泛函理论(Density functional theory, DFT)计算相结合,提出了一种目标驱动方法,发现了一系列未经探索的稳定且无毒的有机-无机混合钙钛矿(Hybrid organic-inorganic perovskite, HOIP),并且构建了新的HOIPs数据库^[13]。Li等以数据驱动的方式高速准确地预测了器件的转换效率(Power conversion efficiency, PCE),并成功地用贝叶斯方法反向预测了相应PSCs每层的厚度、带隙和电子迁移率等基本参数^[14]。Lu和Wei等利用文献中的1072个实际实验数据建立了16个机器学习(Machine learning, ML)模型,从17个PSCs的制造工艺参数中预测PSCs的光伏参数。通过探索影响每个光伏参数的因素,显著提升了器件效率^[15]。Zhan和Zhang团队提出并实现了一种基于正反向框架的机器学习方法,通过建立关键参数与光伏性能之间的关系,实现混合Sn-Pb PSCs光伏性能的优化^[16]。Xu等创建了一个机器学习 workflow来探索伪卤化物阴离子工程,确定了15种潜在的双功能PH阴离子,能够钝化供体和受体,通过实验,发现巯基乙醇酸钠是最有效的钝化剂^[17]。Liu等将知识约束纳入贝叶斯优化,提出一种ML引导的顺序学习框架,在不到100个条件下产生了18.5%的最佳转换效率^[18]。Buonassisi等讨论了阻碍卤化物钙钛矿商业化的一些挑战,并分析了ML工具如何提供帮助^[19]。Zhao等基于多种老化条件下的自动化高通

量实验,分析发现钙钛矿中存在一种稳定性反转的现象,并指出这种现象对钙钛矿光电器件的指导意义^[20]。

考虑到结晶质量是决定钙钛矿效率高低的因素,所以本工作对于钙钛矿太阳能电池制备工艺的优化主要针对光吸收层,从PbI₂过量百分比、退火温度、退火时间以及真空萃取时间四个可控变量来进行。这四个因素是真空刮涂钙钛矿方法影响钙钛矿结晶质量比较重要的参数,将转换效率作为衡量标准,利用贝叶斯优化算法对四个变量进行全局优化,经过多次迭代直到达到符合的预期为止。通过使用随机森林回归对所有实验数据进行拟合,分析其特征重要性。我们发现,PbI₂过量百分比对器件的影响最大,退火温度次之,这为相关研究提供了可借鉴的结果。最终,我们通过简单的工艺参数迭代,实现了效率从22.38%提升至23.56%的进步。

2 实验

2.1 材料

碘化铅(PbI₂, >99.99%)购自于西安亿兰特;碘化甲脒(FAI, >99.99%)购自于Greatcell Solar Materials公司;甲基氯化胺(MACl, >99.99%)购自于Sigma-Aldrich;苯乙基碘化铵(PEAI, >99.5%)、富勒烯C₆₀(>99.5%)均购自于西安宝莱特科技有限公司;N,N-二甲基甲酰胺(DMF, >99.8%)、异丙醇(IPA, >99.5%)、N-甲基吡咯烷酮(NMP, >99.8%)、无水乙醇(EtOH, >99.5%)均购自麦克林。所有药品无需进一步纯化,直接使用。

2.2 制备过程

钙钛矿前驱体溶液:将所需量碘化铅(PbI₂) 1.2+X mol(X与实验条件相关)、碘化甲脒(FAI) 1.08 mol、甲基碘化胺(MAI) 0.12 mol、甲基氯化胺(MACl) 0.18 mol、苯乙基碘化铵(PEAI) 2 mg溶解于1 mL N,N-二甲基甲酰胺(DMF)和氮甲基吡咯烷酮(NMP)(V:V=6:1)混合溶剂中,搅拌至完全溶解。

钙钛矿太阳能电池制备过程:将经过激光刻蚀的2.5 cm×2.5 cm ITO/玻璃基底依次在清洗剂、无水乙醇中超声处理12 min,之后使用氮气吹干ITO/玻璃基底,使用紫外线臭氧设备处理10 min。随后,将1.5 mg/mL的PTAA溶液(溶剂为氯苯)以3 000 r/min旋涂在基板上,100 °C退火10 min。之后将Al₂O₃溶液(用IPA进行1:50稀释)以

3 500 r/min旋涂在PTAA膜层上,100 °C退火10 min。制备钙钛矿薄膜:将刮刀和基底之间的固定间距设置为300 μm,刮刀移动速度设置为3 mm/s,之后使用真空干燥设备(Vacuum concentration drying, VCD)获取中间相,随后放到热板上退火。将样品传送到蒸镀舱内,抽真空至5×10⁻⁴ Pa后以0.02 nm/s进行30 nm的C₆₀薄膜蒸镀,并利用原子层沉积(Atomic layer deposition, ALD)设备沉积厚度约20 nm的SnO₂薄膜,最后通过热蒸发法沉积厚度约为150 nm的Ag电极。

2.3 表征

钙钛矿光伏器件的电流密度-电压(J-V)曲线通过使用数字源表(Keithley 2400)和具有AM 1.5 G光谱的LightSky Tech太阳能模拟器(LSS-55)进行表征。使用标准硅电池将光强调调整为1 000 W/m²,并用孔径面积为0.113 cm²的黑色掩膜放在器件顶部以控制照明区域。光伏器件的外量子效率(External quantum efficiency, EQE)采用光谱响应系统(Enlitech QE-R)测量,该系统以硅参考太阳能电池校准。

2.4 贝叶斯优化

贝叶斯优化是超参数优化最受欢迎的算法之一,它是一种近似逼近的方法,用各种代理函数来拟合超参数与模型评价之间的关系,然后选择有希望的超参数组合进行迭代,最后得出效果最好的超参数组合^[21]。贝叶斯超参数优化主要由以下几个步骤组成:

(1)选择一个代理模型对真实函数 $f(x)$ 建模并定义其先验。由于并不清楚目标函数具体的表达式,需利用已有的知识通过概率代理模型猜测一个目标函数。概率代理模型有高斯过程、随机森林和深度神经网络等,贝叶斯超参数优化通常采用高斯过程作为概率代理函数,高斯回归相关公式如下:

$$f(x) \approx GP(0, k(x, x')), \quad (1)$$

$$p(f(x')|X, Y, x') = N(\mu(x'), \sigma^2(x')), \quad (2)$$

公式(1)中表明 $f(x)$ 是一个高斯过程;公式(2)中 X 是已观测的输入变量, Y 是相应的目标函数值, $\mu(x')$ 是预测的均值, $\sigma^2(x')$ 是相应的方差。

(2)给定一组观测值,通过贝叶斯规则来获取后验。这里第一批观测值使用拉丁超立方采样(Latin hypercube sampling, LHS)来获取。LHS是一种分层采样,可以对变量的分布区间进行高效

采样,主要由分层、采样、乱序三步来进行。不同于随机抽样,这种方法通过最大化地使每一个边缘分布分层,能够保证每一个变量范围的全覆盖,以此来作为贝叶斯优化第一批观测值^[22]。

(3)选择采集函数 $\alpha(x)$ 来确定下一个采样点: $X_i = \operatorname{argmax}_x \alpha(x)$ 。采集函数衡量观测点对拟合 f^* 所产生的影响,并选取影响最大的点作为下一个要观察的点。常见的采集函数主要是概率增量 (Probability of improvement, PI)、期望增量 (Expectation improvement, EI)、置信度上界 (Upper confidence bound, UCB)、信息熵 (Entropy) 等等。在这里采用 EI 作为采集函数去获取下一个实验计划。相较于其他采集函数, EI 采集函数选取期望改进最大的点,并且考虑目标函数预测的不确定性,在平衡探索与利用之间提供了一种权衡:

$$EI(x) = E[\max(f_{i+1}(x) - f(x^*), 0)], \quad (3)$$

其中 x^* 表示历史最佳点。

(4)将采集的数据添加到观测值中,重复执行步骤(2),直到收敛或用尽预算。

2.5 实验设计

为了制备形貌可控、高表面覆盖率和少针孔的高质量钙钛矿薄膜,利用贝叶斯超参数优化法来设计实验。首先对实验所涉及到的四个变量空间进行约束,接着利用 LHS 从预先定义好的变量空间获取第一批实验计划,然后通过刮涂法制备钙钛矿太阳能电池并测试器件性能。利用高斯过程回归模型对输入变量为 PbI_2 过量百分比、退火时间、退火温度、真空萃取时间以及输出变量为器件转换效率的数据进行拟合,核函数选择平滑度参数为 5/2 的 Matern 核函数。Matern 核函数的平滑度参数是一个非负实数,当平滑度参数为 0.5 时,等价于高斯核函数;当平滑度参数趋于无穷大时,Matern 核函数类似于指数核函数。而 Matern 5/2 核函数与其他 Matern 核函数相比,在中间点有更平滑的过渡,在远离中心的地方衰减更快,这种特性使得它在一些空间建模和回归问题中表现较好,同时具有一定的鲁棒性。模型训练结束之后使用 EI 采集函数来获取下一轮的实验计划。我们选择在第一轮生成十六组实验条件,让其尽量覆盖到四个变量空间。之后每轮生成四或五组实验条件。为了确保实验操作的正确性以及模型输入数据的准确性,在控制其他影响实验结果的变量基本不变的情况下,每组条件制备四片尽可能

相同的钙钛矿太阳能电池,选择 PCE 值最高的器件作为 GP 模型的输出变量。每轮实验结束后将数据输入到模型进行训练,直到模型收敛或者达到预期计划要求后停止迭代。

在贝叶斯优化模型迭代结束之后,通过随机森林回归模型来对数据进行训练。随机森林回归模型是以决策树为基学习器的集成学习算法,通过随机抽取样本和特征,建立多棵相互不关联的决策树,通过并行的方式获得预测结果。每棵决策树都能通过抽取的样本和特征得出一个预测结果,通过综合所有树的结果取平均值,得到整个森林的回归预测结果。通过特征重要性评分来分析哪个特征对模型最终预测结果影响更大。

3 结果与讨论

3.1 参数空间

本工作利用贝叶斯优化模型对钙钛矿太阳能电池光吸收层制备所涉及到的四个重要变量进行优化。采用真空萃取方法实现钙钛矿成核,通过精确控制真空萃取的真空度和真空持续时间,可以调整溶液中前驱体溶质的浓度,从而控制晶核的形成速率,提高钙钛矿薄膜质量。为了更好地理解真空萃取时间变量,我们绘制了真空萃取设备 VCD 时间与压强的大致关系图(图 S1)。每个变量的参数空间如表 1 所示。

表 1 工艺参数空间

Tab. 1 The range of process parameter

实验变量	参数空间
PbI_2 过量百分比	(0%, 5%, 0.005%)
退火温度	(100, 150, 5 °C)
退火时间	(10, 20, 1 min)
真空萃取时间	(5, 15, 1 s)

3.2 模型迭代

选出的实验组的工艺参数可以在表 S1 中看到,所有实验组制备器件的转换效率在表 S2 中给出。第一轮实验利用 LHS 根据表中的参数空间均匀设计了十六组实验,经过制备以及表征后得到器件相关性能数据。由于制备过程中一些不可控的噪声,器件表征结果可能会存在一些误差值,这里采用绝对中位差 MAD (Median absolute deviation) 来检测异常值。MAD 法是一种基于中值绝对偏差的去噪方法。首先分别计算每个条件所制备器件的异常值,接着对存在异常值的数据进行

去除之后再绘制中位数图。根据器件最终转换效率,我们设置5,10,15,20四个分界值,并采用不同颜色对应分界值去绘制每组实验所制备器件的中位数图(图1)。可以发现实验组5~8以及14的

转换效率相对较低,进一步研究具体实验条件可以发现,对于退火温度过高、真空萃取时间较短、 PbI_2 过量百分比低这三个可能影响薄膜质量的因素,在这几组实验中都有涉及到三者之一。

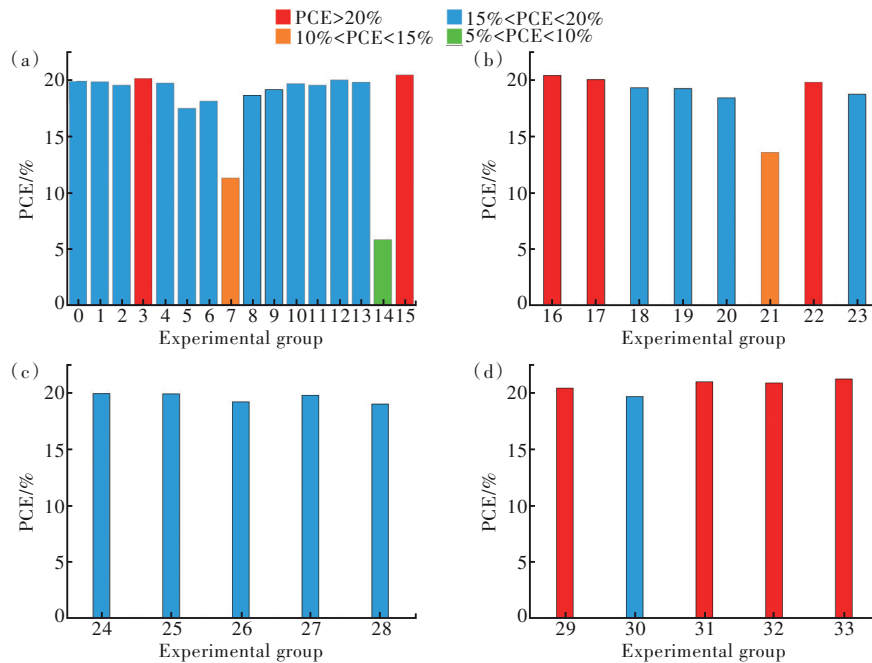


图1 (a)第一轮制备器件效率中位数;(b)第二、三轮制备器件效率中位数;(c)第四轮制备器件效率中位数;(d)第五轮制备器件效率中位数

Fig.1 Median device efficiency of preparation for the first round(a), the second and third rounds(b), the fourth round(c), the fifth round(d)

将第一批实验的数据处理之后作为贝叶斯优化的先验数据,对数据进行拟合后利用EI采集函数进行下一批实验计划的选择。观察第二批和第三批器件的PCE中位数图,可以看出图中表现较好的实验组为16,17,22。进一步观察发现,这几个条件的碘化铅过量百分比其实相差较小,结合第一批实验所得结论,考虑到可能碘化铅含量、退火温度对实验结果影响偏大,所以为了让采集函数进行更精确的探索,将碘化铅过量百分比的参数空间设置为2%~5%,步长设置为0.0025%,同时将退火温度参数空间设置为100~130℃,步长为5℃。

同样,将前几批实验所得数据输入到模型来选择下一批实验计划,器件制备完成后进行表征。通过观察第四批、第五批器件的PCE中位数图可以发现,这两批实验相较于之前来说整体数据中位数相差并没有很大,可以推断出模型拟合效果不错,而且采集函数选取的点也会更有利于器件转换效率的提升。尤其是对于第五批实验,即实验组29~33。五个实验组中有四个实验组的PCE

中位数都达到20%以上,最差的条件30也为19.7%,最好的条件33的PCE中位数为21.25%,对比碘化铅不过量所制备的器件PCE中位数提高了3.59%。根据第一轮部分实验与最后一轮实验结果绘制了器件性能箱线图,如图2所示,图中1~5为第一轮实验,29~33为第五轮实验。由于这里对器件性能相关参数进行了约束,选取短路电流(J_{sc})为23 mA/cm²至30 mA/cm²之间、开路电压(V_{oc})大于0.7 V、填充因子在0~1内,所以有些实验组显示数据较少。可以很明显地看出,经过四轮工艺条件优化后,器件的开路电压和填充因子都有了明显的提升,相应的效率也会提升。通过实验制备了最高转换效率的器件,器件整体结构可在图3(a)中观察到,其转换效率为23.56%。我们对该组工艺条件进行了重复实验,图S2中给出了相应的结果。如图3(c)所示,我们测试了相应器件的EQE,获得了24.61 mA/cm²的积分电流,和I-V测试(图3(b))所得 J_{sc} 相差1.19 A/cm²,在5%以内。通过给器件施加0.95 V的正向偏压持续测量电流密度,并计算获得相应的输出效率

率,将其绘制成与测量时间的函数图像(图3(d))。未封装器件在150 s内表现出高度稳定的22.4%的输出效率。我们另外测试了XRD(图S3),结

果显示薄膜的结晶以(110)取向为主,且没有杂峰存在,证明我们的工艺可以获得高质量的钙钛矿薄膜。

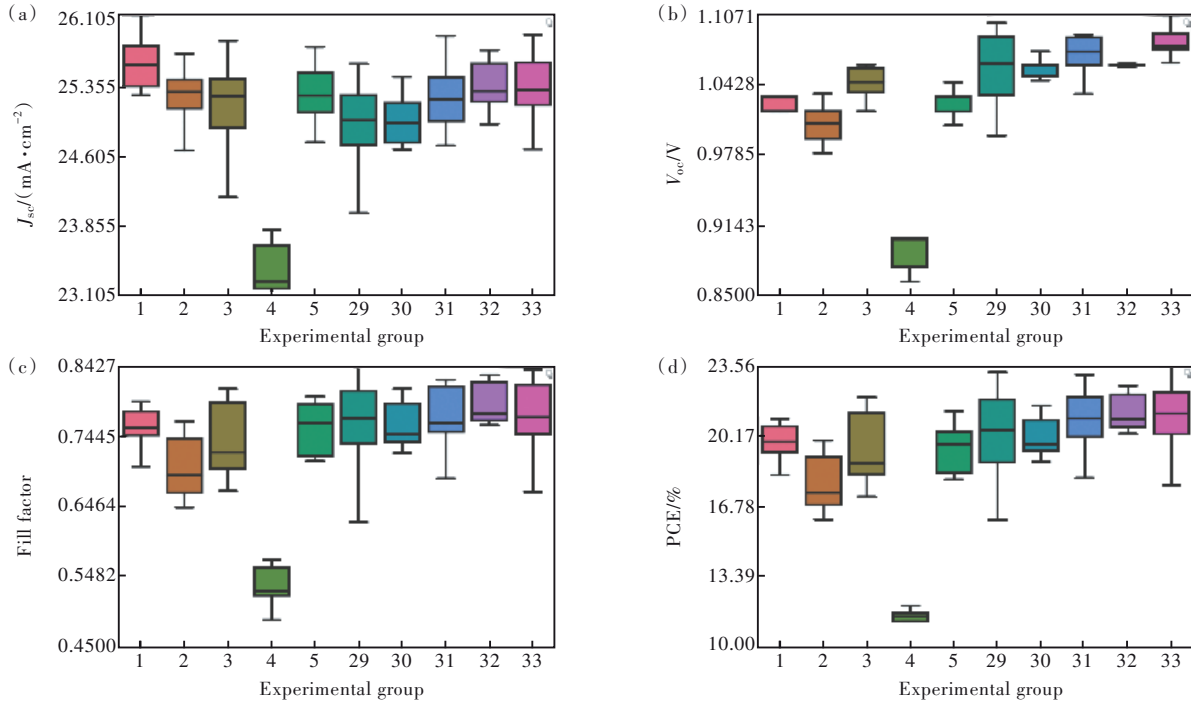


图2 第一轮与第五轮所制备器件性能对比。(a)短路电流;(b)开路电压;(c)填充因子;(d)转换效率

Fig.2 Comparison of the performance parameters of the devices prepared in the first and fifth rounds. (a) Short-circuit current. (b) Open-circuit voltage. (c) Fill factor. (d) Conversion efficiency

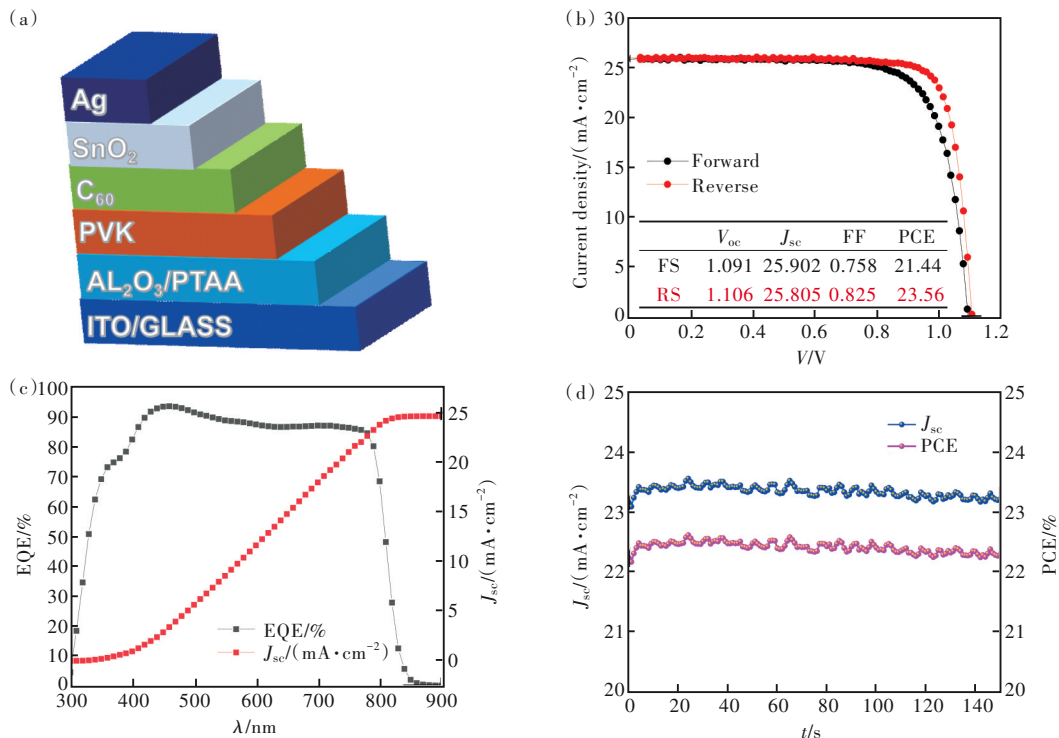


图3 (a)器件整体结构图;(b)冠军器件的I-V曲线;(c)EQE光谱和相应的电流密度;(d)稳态功率输出测试

Fig.3 (a)The overall structure diagram of the device. (b)I-V curves of the champion device. (c)EQE spectra and corresponding integrated current densities test. (d)Steady-state power output test

3.3 特征重要性评估

迭代实验结束之后,将所有批次实验数据进行整理,找出每个实验组的最优器件作为随机森林回归模型的输入,借助 sklearn 库的 RandomForestRegressor 包来构建模型,在模型训练完成后计算特征重要性得分。本实验所涉及到的四个变量对于器件效率的重要性程度如图4所示。可以看到影响最大的是 PbI_2 过量百分比,达到 0.621,这与实验过程中观察到的结果相一致。其次是退火温度 0.206,相比较而言,影响较小的是退火时间与真空萃取时间。少量 PbI_2 被证明有利于提高钙钛矿薄膜的质量^[23-25],而退火温度和时间直接影响了 PbI_2 残余量的多少,因此 PbI_2 含量的重要程度高于退火温度和时间。抽真空时间影响

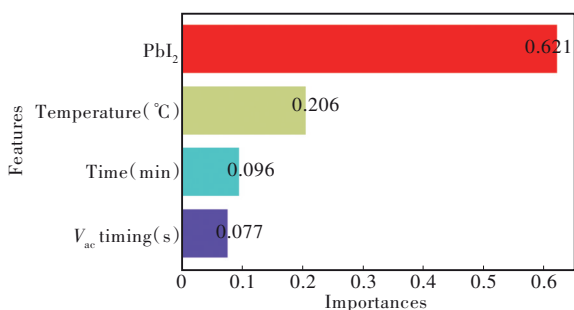


图4 工艺参数特征重要性

Fig.4 The importance of process parameter characteristics

最小,说明在我们设置的参数范围内,它的工艺窗口是比较宽的。由此可以证明使用贝叶斯模型来对工艺参数进行优化是合理的,模型可以学习到不同工艺参数的内在关联进而推荐全局最优的工艺参数组合。

4 结 论

本工作主要通过机器学习模型来对钙钛矿太阳能电池制备工艺进行优化。利用贝叶斯优化模型对制备过程中所涉及到的碘化铅过量百分比、退火温度、退火时间以及真空萃取时间四个特征参数空间进行优化,经过五轮迭代和参数调整,制备出 23.56% 的最高效率器件,在最后一轮选出的五组实验中有四组出现转换效率超过 23% 的器件。通过随机森林回归模型对实验数据进行建模,分析了实验中四个工艺参数的重要性程度。本工作利用机器学习模型辅助制备钙钛矿太阳能电池,显著缩短了研发进程,相较于实验试错法,在成本更低、时间更短的基础上制备了高效率的钙钛矿太阳能电池。

本文补充文件和专家审稿意见及作者回复内容的下载地址:<http://cjl.lightpublishing.cn/thesisDetails#10.37188/CJL.20230309>。

参 考 文 献:

- [1] KOJIMA A, TESHIMA K, SHIRAI Y, *et al.* Organometal halide perovskites as visible-light sensitizers for photovoltaic cells [J]. *J. Am. Chem. Soc.*, 2009, 131(17): 6050-6051.
- [2] LIU C, CHENG Y B, GE Z Y. Understanding of perovskite crystal growth and film formation in scalable deposition processes [J]. *Chem. Soc. Rev.*, 2020, 49(6): 1653-1687.
- [3] RAMPRASAD R, BATRA R, PILANIA G, *et al.* Machine learning in materials informatics: recent applications and prospects [J]. *npj Comput. Mater.*, 2017, 3(1): 54.
- [4] DOGAN A, BIRANT D. Machine learning and data mining in manufacturing [J]. *Expert Syst. Appl.*, 2021, 166: 114060.
- [5] SELVARATNAM B, KOODALI R T. Machine learning in experimental materials chemistry [J]. *Catal. Today.*, 2021, 371: 77-84.
- [6] SNOEK J, LAROCHELLE H, ADAMS R P. Practical Bayesian optimization of machine learning algorithms [C]. *Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe*, 2012: 2951-2959.
- [7] LI F, PENG X Q, WANG Z, *et al.* Machine learning(ML)-assisted design and fabrication for solar cells [J]. *Energy Environ. Mater.*, 2019, 2(4): 280-291.
- [8] YAN W S, LIU Y M, ZANG Y, *et al.* Machine learning enabled development of unexplored perovskite solar cells with high efficiency [J]. *Nano Energy*, 2022, 99: 107394.
- [9] ZHANG Y, XU X J. Machine learning lattice constants for cubic perovskite A_2XY_6 compounds [J]. *J. Solid State Chem.*, 2020, 291: 121558.

- [10] LI L, TAO Q L, XU P C, *et al.* Studies on the regularity of perovskite formation *via* machine learning [J]. *Comput. Mater. Sci.*, 2021, 199: 110712.
- [11] SUN S J, HARTONO N T P, REN Z D, *et al.* Accelerated development of perovskite-inspired materials *via* high-throughput synthesis and machine-learning diagnosis [J]. *Joule*, 2019, 3(6): 1437-1451.
- [12] VAKHARIA V, CASTELLI I E, BHAVSAR K, *et al.* Bandgap prediction of metal halide perovskites using regression machine learning models [J]. *Phys. Lett. A*, 2022, 422: 127800.
- [13] WU T M, WANG J. Global discovery of stable and non-toxic hybrid organic-inorganic perovskites for photovoltaic systems by combining machine learning method with first principle calculations [J]. *Nano Energy*, 2019, 66: 104070.
- [14] LI W H, HU J H, CHEN Z X, *et al.* Performance prediction and optimization of perovskite solar cells based on the Bayesian approach [J]. *Sol. Energy*, 2023, 262: 111853.
- [15] LU Y, WEI D, LIU W, *et al.* Predicting the device performance of the perovskite solar cells from the experimental parameters through machine learning of existing experimental results [J]. *J. Energy Chem.*, 2023, 77: 200-208.
- [16] CAI X, LIU F C, YU A R, *et al.* Data-driven design of high-performance $\text{MASn}_x\text{Pb}_{1-x}\text{I}_3$ perovskite materials by machine learning and experimental realization [J]. *Light Sci. Appl.*, 2022, 11(1): 234.
- [17] XU J, CHEN H, GRATER L, *et al.* Anion optimization for bifunctional surface passivation in perovskite solar cells [J]. *Nat. Mater.*, 2023, 22(12): 1507-1514.
- [18] LIU Z, ROLSTON N, FLICK A C, *et al.* Machine learning with knowledge constraints for process optimization of open-air perovskite solar cell manufacturing [J]. *Joule*, 2022, 6(4): 834-849.
- [19] KUMAR R E, TIHONEN A, SUN S J, *et al.* Opportunities for machine learning to accelerate halide-perovskite commercialization and scale-up [J]. *Matter*, 2022, 5(5): 1353-1366.
- [20] ZHAO Y C, ZHANG J Y, XU Z W, *et al.* Discovery of temperature-induced stability reversal in perovskites using high-throughput robotic learning [J]. *Nat. Commun.*, 2021, 12(1): 2191.
- [21] SHAHRIARI B, SWERSKY K, WANG Z Y, *et al.* Taking the human out of the loop: a review of Bayesian optimization [J]. *Proc. IEEE*, 2016, 104(1): 148-175.
- [22] LOH W L. On Latin hypercube sampling [J]. *Ann. Statist.*, 1996, 24(5): 2058-2080.
- [23] WANG L L, MCCLEESE C, KOVALSKY A, *et al.* Femtosecond time-resolved transient absorption spectroscopy of $\text{CH}_3\text{NH}_3\text{PbI}_3$ perovskite films: evidence for passivation effect of PbI_2 [J]. *J. Am. Chem. Soc.*, 2014, 136(35): 12205-12208.
- [24] CAO D H, STOUMPOS C C, MALLIAKAS C D, *et al.* Remnant PbI_2 , an unforeseen necessity in high-efficiency hybrid perovskite-based solar cells? [J]. *APL Mater.*, 2014, 2(9): 091101.
- [25] LEE Y H, LUO J S, HUMPHRY-BAKER R, *et al.* Unraveling the reasons for efficiency loss in perovskite solar cells [J]. *Adv. Funct. Mater.*, 2015, 25(25): 3925-3933.



弓箭(1997-),男,山西晋中人,硕士研究生,2019年于吕梁学院获得学士学位,主要从事人工智能与新型钙钛矿太阳能电池方面的相关研究。
E-mail: 754138216@qq.com



刘冲(1989-),男,河北保定人,博士,副研究员,2020年于暨南大学获得博士学位,主要从事新型钙钛矿太阳能电池研发(涉及大面积涂布和组件开发等领域)。
E-mail: chongliu@jnu.edu.cn



李阳(1969-),男,辽宁鞍山人,博士,教授,2000年于南开大学获得博士学位,主要从事光电材料与器件、人工智能等方面的研究。
E-mail: insidesun51@163.com