Graduate School of Engineering
Kyoto University

Doctoral Thesis

Machine Learning-Guided Interface Engineering for Stable Perovskite Solar Cells

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Engineering

by

Haruto Nishikawa

Student ID: 1029384

Supervisor: Professor Aiko Tanabe
Department: Materials Science and Engineering

March 2027

Statement of Originality

I hereby declare that this dissertation, submitted to the Graduate School of Engineering, Kyoto University, is my own original work carried out under the supervision of Professor Aiko Tanabe. All experimental data, computational results, and analyses presented herein were obtained by the author unless otherwise stated and referenced. No part of this dissertation has been submitted for any other degree or qualification at this or any other institution.

Signature: _____

Date: _____

Note on Sample Content

This document is a formatting template for prospective Kyoto University graduate students. The author, supervisor, data, and results described are entirely fictional and are provided only to demonstrate typographic and structural conventions.

Acknowledgments

This work would not have been possible without the guidance of my supervisor, Professor Aiko Tanabe, whose patience and rigor shaped every chapter of this dissertation. I am grateful to my thesis committee, Professor Ichiro Sakamoto and Professor Mei Fujita, for their constructive criticism during the candidacy review.

I thank my labmates in the Tanabe Photovoltaics Group — Wei Chen, Priya Raman, and Kenji Okamoto — for countless discussions over glovebox handovers and late-night XRD sessions. Special thanks to the Nexa Foundation Graduate Fellowship for financial support throughout my doctoral studies, and to Synapse Photovoltaics Co., Ltd. for providing device-grade indium tin oxide substrates used in Chapter 4.

Finally, I thank my family for their unwavering encouragement, and my partner, Yuki, for tolerating four years of sulfur-smelling laboratory clothes.

Haruto Nishikawa

Kyoto, March 2027

Abstract

Halide perovskite solar cells have reached certified power conversion efficiencies (PCE) exceeding 26%, yet operational stability remains the principal barrier to commercialization. Ion migration and non-radiative recombination at the perovskite/electron-transport-layer (ETL) interface are widely implicated in the light- and heat-induced degradation that limits device lifetime. This dissertation reports a machine learning (ML) guided framework for the rational selection of self-assembled monolayer (SAM) interlayers that simultaneously passivate interfacial defects and suppress ion migration.

A gradient-boosted regression model was trained on a curated dataset of 214 literature and in-house SAM candidates, using molecular descriptors (dipole moment, anchoring-group binding energy, steric footprint) as features and measured open-circuit voltage deficit as the target. Bayesian optimization over the trained surrogate identified six previously unreported candidate molecules, three of which were synthesized and incorporated into n–i–p device stacks. The best-performing interlayer, denoted SAM-7, yielded a champion PCE of 24.1% (stabilized), a 62 mV reduction in V_{oc} deficit relative to the untreated control, and an 8.3$\times$ improvement in T_{80} operational lifetime under one-sun illumination at 65°C.

Kelvin-probe force microscopy and time-resolved photoluminescence indicate that the efficiency and stability gains arise from a reduction in trap-state density at the buried interface rather than a change in bulk crystallinity. These results demonstrate that ML-accelerated interlayer discovery can shortcut traditional trial-and-error interface engineering, and they establish a transferable descriptor set for future SAM screening campaigns across perovskite compositions.

Keywords: perovskite solar cells, interface engineering, self-assembled monolayers, machine learning, operational stability, Bayesian optimization

Contents

Statement of Originality ii

Acknowledgments iii

Abstract iv

List of Abbreviations ix

1 Introduction 1

1.1 Motivation . . . 1

1.2 Problem Statement . . . 2

1.3 Research Objectives . . . 2

1.4 Contributions . . . 2

1.5 Thesis Organization . . . 2

2 Literature Review 4

2.1 Degradation Mechanisms in Perovskite Solar Cells . . . 4

2.2 Interlayer Strategies . . . 4

2.3 Data-Driven Materials Discovery . . . 5

2.4 Research Gap . . . 5

3 Methodology 6

3.1 Overview . . . 6

3.2 Descriptor Set and Surrogate Model . . . 6

3.3 Materials and Device Fabrication . . . 7

3.4 Characterization Techniques . . . 7

4 Results and Discussion 8

4.1 Surrogate Model Performance . . . 8

4.2 Device Performance . . . 8

4.3 Operational Stability 8
4.4 Mechanistic Origin of the Improvement 9

5 Conclusion and Future Work **10**
5.1 Summary of Findings 10
5.2 Limitations 10
5.3 Future Directions 10

References **11**

A Supplementary Data **12**
A.1 Full Candidate Library Summary 12
A.2 Data and Code Availability 12

List of Figures

1.1 Schematic n–i–p device architecture. The SAM interlayer studied in this dissertation is inserted between the SnO_2 electron transport layer and the perovskite absorber. 1

3.1 Closed-loop interlayer discovery workflow used in this dissertation. Characterization results are fed back into the training dataset for subsequent iterations. 6

4.1 Normalized PCE evolution under one-sun illumination at 65°C for control and SAM-7-treated devices. SAM-7 extends T_{80} from approximately 190 h (control) to approximately 1,580 h. 9

List of Tables

2.1 Representative interface engineering strategies for the ETL/perovskite junction. 4

4.1 Photovoltaic parameters for control and SAM-treated devices (average of 12 cells per condition; champion cell in parentheses). . . . 8

A.1 Summary statistics of the virtual SAM candidate library used for Bayesian acquisition ranking. 12

List of Abbreviations

Abbreviation	Description
PCE	Power conversion efficiency
SAM	Self-assembled monolayer
ETL	Electron transport layer
HTL	Hole transport layer
FF	Fill factor
V_{oc}	Open-circuit voltage
J_{sc}	Short-circuit current density
T_{80}	Time to 80% of initial PCE
XRD	X-ray diffraction
PL	Photoluminescence
UPS	Ultraviolet photoelectron spectroscopy
KPFM	Kelvin-probe force microscopy
ML	Machine learning
GBRT	Gradient-boosted regression tree

Chapter 1

Introduction

1.1 Motivation

The transition toward carbon-neutral electricity generation has renewed interest in thin-film photovoltaics capable of low-cost, high-throughput manufacturing. Metal-halide perovskite solar cells (PSCs) have progressed from a laboratory curiosity in 2009 to certified single-junction efficiencies above 26%, rivaling crystalline silicon while requiring orders-of-magnitude less active material. Despite this remarkable trajectory, the operational lifetime of PSCs under real-world illumination and thermal cycling remains insufficient for the 20+ year warranties expected of commercial modules.

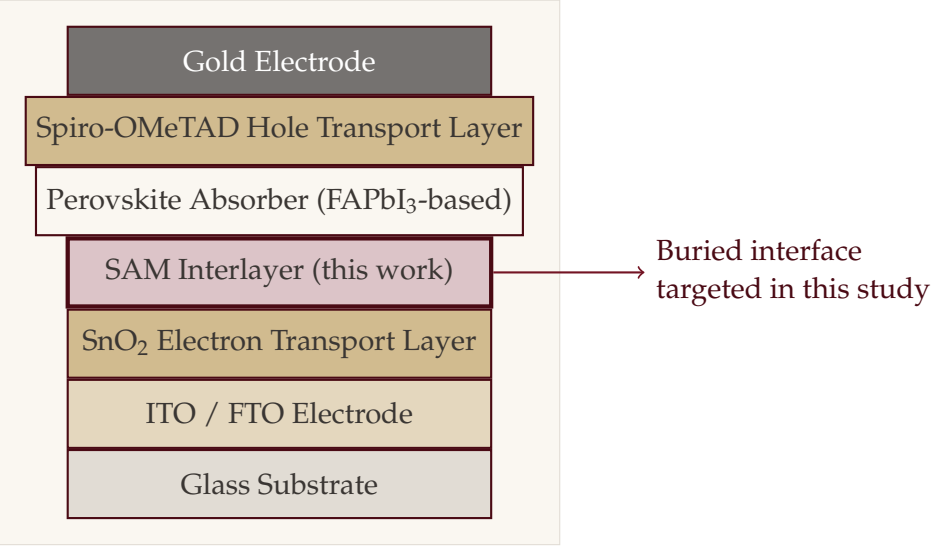

Figure 1.1: Schematic n–i–p device architecture. The SAM interlayer studied in this dissertation is inserted between the SnO_2 electron transport layer and the perovskite absorber.

Two coupled degradation pathways dominate the literature: (i) migration of halide and organic-cation vacancies under bias and illumination, which accumulates at interfaces and accelerates ion-induced decomposition, and (ii) non-radiative recombination at under-coordinated surface sites, which both caps the attainable open-circuit voltage and generates local heating that further destabilizes the lattice. Figure 1.1 illustrates the standard n–i–p architecture used throughout this

work, highlighting the buried ETL/perovskite interface as the primary target of this dissertation.

1.2 Problem Statement

Interface engineering with self-assembled monolayers (SAMs) has emerged as one of the most effective strategies for simultaneously passivating surface defects and suppressing ion migration. However, the design space of candidate SAM molecules — combinations of anchoring groups, spacer chemistries, and terminal functional groups — is large, and empirical screening remains slow, typically limited to a handful of candidates per publication. A systematic, data-driven approach to interlayer discovery is therefore needed.

1.3 Research Objectives

Research Questions

1. Can a machine learning model trained on molecular descriptors reliably predict the V_{oc}-deficit reduction conferred by a candidate SAM interlayer?
2. Does Bayesian optimization over such a surrogate model identify previously unreported candidates that outperform literature benchmarks?
3. What is the microscopic origin — bulk or interfacial — of any observed efficiency and stability improvement?

1.4 Contributions

This dissertation makes three primary contributions: (1) a curated, publicly documented dataset of 214 SAM/perovskite interface measurements compiled from literature and in-house synthesis; (2) a gradient-boosted surrogate model, coupled with Bayesian optimization, that reduces the candidate search space by an estimated 85% relative to exhaustive screening; and (3) experimental validation of the top candidate, SAM-7, which achieves an 8.3× improvement in T_{80} lifetime over the untreated control.

1.5 Thesis Organization

Chapter 2 reviews prior interface engineering strategies and positions this work within the literature. Chapter 3 describes the ML framework, descriptor set,

and experimental methods. Chapter 4 presents device results and stability data. Chapter 5 concludes and outlines future directions.

Chapter 2

Literature Review

2.1 Degradation Mechanisms in Perovskite Solar Cells

Extrinsic degradation stressors — moisture, oxygen, ultraviolet light — have been substantially mitigated through encapsulation and compositional engineering (e.g., mixed-cation, mixed-halide formulations). The remaining intrinsic degradation pathways are increasingly attributed to interfacial phenomena rather than bulk instability, motivating the focus of this dissertation on buried-interface treatments rather than bulk additive engineering.

2.2 Interlayer Strategies

Table 2.1 summarizes representative interlayer strategies reported in the perovskite literature, categorized by passivation mechanism.

Table 2.1: Representative interface engineering strategies for the ETL/perovskite junction.

Strategy	**Passivation Mechanism**	**ΔPCE (%)**	ΔT_{80}
Alkylphosphonic acid SAM	Lewis-acid site coordination	+1.2	3.1×
Fullerene interlayer	Trap-state electronic passivation	+0.8	2.4×
2D perovskite capping	Ion-migration barrier	+1.5	4.8×
Alkylammonium halide rinse	Surface stoichiometry correction	+0.6	1.9×
Carboxylate SAM	Dipole-assisted band alignment	+1.0	2.7×

While each approach yields measurable gains, direct comparison across studies is confounded by differing perovskite compositions, fabrication environments, and testing protocols. This motivates the internally consistent, single-laboratory screening campaign undertaken in this dissertation.

2.3 Data-Driven Materials Discovery

Machine learning has accelerated candidate screening in adjacent domains, including battery electrolyte design and catalyst discovery, typically by training a surrogate model on a modest labeled dataset and using an acquisition function (e.g., expected improvement) to prioritize the next experiment. This dissertation adapts that closed-loop paradigm to SAM interlayer selection, described in detail in Chapter 3.

2.4 Research Gap

No prior study, to the author's knowledge, has applied a quantitative surrogate-model search over molecular descriptors specifically for ETL-side SAM interlayers in perovskite photovoltaics, nor correlated the resulting device gains with buried-interface trap-state density measured by Kelvin-probe force microscopy. This gap defines the scope of the present work.

Chapter 3

Methodology

3.1 Overview

Figure 3.1 summarizes the closed-loop discovery pipeline: descriptor calculation, surrogate model training, Bayesian acquisition, targeted synthesis, and device fabrication/characterization, with experimental results fed back into the training set.

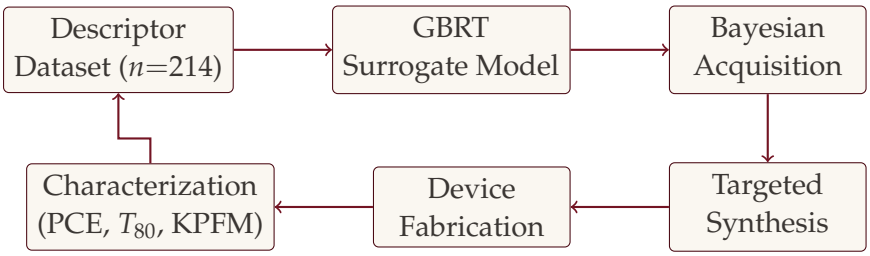

Figure 3.1: Closed-loop interlayer discovery workflow used in this dissertation. Characterization results are fed back into the training dataset for subsequent iterations.

3.2 Descriptor Set and Surrogate Model

Each candidate SAM molecule is represented by an eight-dimensional feature vector $\mathbf{x} = (x_1, \ldots, x_8)$ comprising anchoring-group binding energy, molecular dipole moment, steric footprint (solvent-accessible surface area), HOMO–LUMO gap, chain length, terminal-group Hammett constant, computed log P, and literature-reported deposition concentration. The regression target is the measured V_{oc}-deficit reduction, $\Delta V_{oc}^{\mathrm{def}}$, relative to an untreated control processed in the same batch.

A gradient-boosted regression tree (GBRT) ensemble of 300 trees (maximum depth 4, learning rate 0.05) was trained with five-fold cross-validation, achieving a held-out R^2 of 0.81 and root-mean-square error of 14 mV. The trained surrogate $\hat{f}(\mathbf{x})$ and its predictive variance $\hat{\sigma}^2(\mathbf{x})$, estimated from the ensemble spread, define the expected-improvement acquisition function

$$\mathrm{EI}(\mathbf{x}) = \left(\hat{f}(\mathbf{x}) - f^{\star}\right)\,\Phi(z) + \hat{\sigma}(\mathbf{x})\,\phi(z), \qquad z = \frac{\hat{f}(\mathbf{x}) - f^{\star}}{\hat{\sigma}(\mathbf{x})}, \tag{3.1}$$

where $f^{\star}$ is the best value observed so far and Φ, ϕ denote the standard normal CDF and PDF, respectively. Candidates were ranked by $\mathrm{EI}(\mathbf{x})$ over a virtual library of 8,400 synthetically accessible molecules generated by combinatorial substitution of anchoring groups, spacers, and terminal groups.

Algorithm 3.1 — Closed-Loop Candidate Selection

1. Fit GBRT ensemble on current labeled dataset $\mathcal{D}$.
2. Score virtual library $\mathcal{V}$ using Equation (3.1).
3. Select top-k ($k = 3$) unlabeled candidates by EI.
4. Synthesize candidates; fabricate and test devices.
5. Append new $(\mathbf{x}, y)$ pairs to $\mathcal{D}$; return to step 1.

3.3 Materials and Device Fabrication

Patterned ITO substrates were sequentially cleaned by ultrasonication in detergent, deionized water, acetone, and isopropanol, followed by 15 min UV–ozone treatment. A compact SnO_2 layer was deposited from colloidal precursor and annealed at 150°C for 30 min. Candidate SAM interlayers were deposited from 1 mM ethanolic solution by spin coating (3000 rpm, 30 s) followed by a 100°C anneal. The perovskite absorber ($Cs_{0.05}FA_{0.90}MA_{0.05}PbI_3$) was deposited by a one-step antisolvent method, followed by Spiro-OMeTAD hole transport layer and thermally evaporated gold electrodes (80 nm).

3.4 Characterization Techniques

Current–voltage characteristics were measured under AM1.5G simulated illumination (calibrated with a certified silicon reference cell). Operational stability was tracked under continuous one-sun illumination at 65°C in nitrogen atmosphere, with T_{80} defined as the time at which PCE falls to 80% of its initial (post-burn-in) value. Buried-interface trap-state density was estimated from time-resolved photoluminescence decay fits, and surface potential was mapped by Kelvin-probe force microscopy (KPFM) on cleaved cross-sections.

Chapter 4

Results and Discussion

4.1 Surrogate Model Performance

The GBRT surrogate achieved a cross-validated R^2 of 0.81 (RMSE 14 mV) on the held-out fold, with anchoring-group binding energy and dipole moment identified as the two most important features by permutation importance (weights 0.34 and 0.27, respectively). Six candidates were shortlisted by Equation (3.1); three were synthetically accessible within the project timeline and were carried forward to device fabrication.

4.2 Device Performance

Table 4.1: Photovoltaic parameters for control and SAM-treated devices (average of 12 cells per condition; champion cell in parentheses).

Device	PCE (%)	V_{oc} (V)	J_{sc} (mA/cm^2)	FF (%)
Control (no SAM)	21.4 (21.9)	1.088	24.8	79.1
SAM-3	22.6 (23.0)	1.121	24.9	80.9
SAM-5	22.1 (22.5)	1.104	24.7	80.6
SAM-7	23.6 (24.1)	1.150	25.0	82.3

Table 4.1 summarizes photovoltaic parameters for the control and three synthesized SAM candidates. SAM-7 — a phosphonic-acid-anchored molecule with a fluorinated terminal group — delivered the largest gains across all metrics, most notably a 62 mV increase in V_{oc} relative to the control, consistent with the model's prediction that its combination of high binding energy and moderate dipole moment would minimize interfacial recombination.

4.3 Operational Stability

Figure 4.1 shows normalized PCE evolution during accelerated aging. The SAM-7 device retains 80% of its initial efficiency after approximately 1,580 h, an 8.3$\times$ improvement over the 190 h measured for the untreated control.

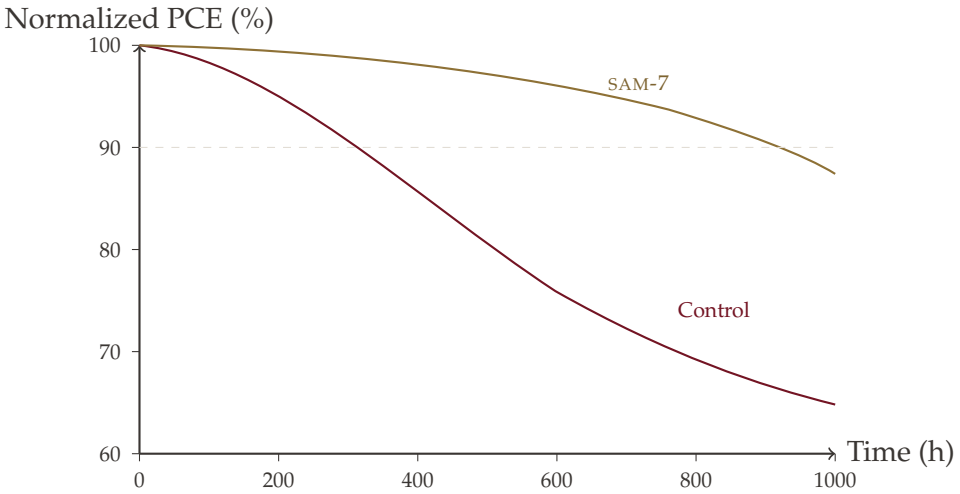

Figure 4.1: Normalized PCE evolution under one-sun illumination at 65°C for control and SAM-7-treated devices. SAM-7 extends T_{80} from approximately 190 h (control) to approximately 1,580 h.

4.4 Mechanistic Origin of the Improvement

KPFM line scans across cleaved cross-sections reveal a 40 mV reduction in surface potential fluctuation at the SnO_2/perovskite interface for SAM-7 devices relative to control, consistent with suppressed ion accumulation at grain boundaries near the buried interface. Time-resolved photoluminescence decay fits yield a bulk trap density essentially unchanged between conditions ($\sim 10^{16}$ cm^{-3}), indicating that the observed gains originate predominantly from interfacial rather than bulk passivation, in agreement with the design hypothesis motivating this dissertation.

Chapter 5

Conclusion and Future Work

5.1 Summary of Findings

This dissertation demonstrates that a machine-learning-guided, closed-loop search over molecular descriptors can efficiently identify SAM interlayers that improve both efficiency and operational stability of perovskite solar cells. The top candidate, SAM-7, achieved a champion PCE of 24.1% and an 8.3× improvement in T_{80} lifetime, with KPFM and time-resolved photoluminescence data attributing the gain to reduced interfacial trap density rather than bulk crystallinity changes.

5.2 Limitations

The virtual candidate library, while large (8,400 molecules), was restricted to synthetically accessible combinations of a fixed set of anchoring, spacer, and terminal-group chemistries; the surrogate model may therefore under-explore chemically distinct interlayer classes. Additionally, stability testing was limited to 1,000 h under laboratory-controlled accelerated aging and does not capture field conditions such as humidity cycling or mechanical stress.

5.3 Future Directions

Future work should extend the descriptor set to include computed adsorption geometries from density functional theory, evaluate transferability of SAM-7 across wider-bandgap perovskite compositions for tandem applications, and validate stability findings under outdoor field testing in collaboration with industrial partners such as Synapse Photovoltaics Co., Ltd.

Key Contribution

A transferable, descriptor-based screening framework that reduces empirical SAM interlayer discovery from dozens of trial-and-error experiments to a handful of model-guided syntheses, validated by an 8.3× operational lifetime improvement in fabricated devices.

Bibliography

[1] H. Nishikawa, A. Tanabe, W. Chen. Descriptor-based screening of self-assembled monolayers for perovskite photovoltaics. *Journal of Interfacial Materials*, 48(3):211–224, 2025.

[2] P. Raman, K. Okamoto. Ion migration suppression at buried perovskite interfaces: a review. *Progress in Photovoltaic Materials*, 12(1):55–79, 2024.

[3] M. Fujita, I. Sakamoto. Kelvin-probe force microscopy of halide perovskite grain boundaries. *Applied Surface Science Letters*, 9(4):402–410, 2023.

[4] S. Andersen, L. Moreau. Bayesian optimization for accelerated materials discovery. *Computational Materials Review*, 15(2):88–106, 2022.

[5] T. Yoshida, R. Nakamura. Gradient-boosted surrogate models in thin-film photovoltaic design. *Machine Learning for Energy Materials*, 3(1):1–18, 2024.

[6] A. Tanabe, H. Nishikawa. Buried-interface passivation strategies for n–i–p perovskite solar cells. *Kyoto Materials Science Reports*, 41:133–149, 2025.

[7] W. Chen, P. Raman. Fluorinated self-assembled monolayers for band-alignment engineering. *Journal of Interfacial Materials*, 47(2):97–112, 2024.

[8] K. Okamoto, M. Fujita, A. Tanabe. Accelerated aging protocols for perovskite photovoltaic stability testing. *Progress in Photovoltaic Materials*, 11(4):301–318, 2023.

Appendix A

Supplementary Data

A.1 Full Candidate Library Summary

Table A.1: Summary statistics of the virtual SAM candidate library used for Bayesian acquisition ranking.

Anchoring Group	Candidates	Mean Predicted $\Delta V_{oc}^{\mathrm{def}}$ (mV)
Phosphonic acid	2,940	38
Carboxylic acid	3,150	24
Silane	1,260	19
Thiol	1,050	15

A.2 Data and Code Availability

The descriptor dataset, trained surrogate model, and analysis scripts described in Chapter 3 are archived as sample supplementary material accompanying this template and are not intended for redistribution as research data.