Approaching Novel Perovskites Photovoltaic Devices through Machine Learning and Interfacial Engineering

by

Ruiqi Zhang

B.S., Nanoengineering, University of California, San Diego (2021)

Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2023

Authored by: Ruiqi Zhang
Department of Electrical Engineering and Computer Science
August 31, 2023

Certified by: Vladimir Bulović
Fariborz Maseeh Professor in Emerging Technologies
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by: Leslie A Kolodziejski
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Students

Abstract

Organic metal halide perovskites have shown plenty of extraordinary optoelectronic properties which make them good candidates for various photovoltaic applications [1-5]. The fascinating optoelectronic properties of perovskite largely take credit to their low exciton binding energy, strong light absorption coefficient, relatively long carrier diffusion length, and carrier recombination lifetime [6-9]. However, even with an increasing number of studies carried out, perovskite solar cell is still facing plenty of challenges towards commercialization. Two main challenges towards large-area commercialization include first the harsh fabrication environment and cost of large-area coating; and second the redundant fabrication process with a huge labor force impelled. In this thesis study, an intermedia thin film layer tris(4-carbazoyl-9-ylphenyl)amine (TcTa) with a thickness of 3 nm is discovered in a large-area compatible perovskite solar cell structure ITO/SnO_2/$(MAFACs)_1Pb(IBrCl)_3$/PV2000/TcTa/Au that reaches a power conversion efficiency above 14%. The TcTa intermediate film is compatible with substituting gold top electrodes and preventing sputter damage while maintaining a similar solar cell performance (etc. sputtered Ni). In addition, a machine learning algorithm is developed to predict the solar cell current-voltage properties only based on the film stack optical properties before the solar cell is fabricated. The algorithm is developed and tested based on the 3D/2D perovskite solar cell structure [10] with resulting in an average prediction regression loss below 5% and a best prediction accuracy above 99%. Multiple different machine learning algorithm is also carried out to analyze the prediction results and learning weights for the model.

Acknowledgement

I would like to first thank my Thesis Supervisor and my research advisor Professor Vladimir Bulović. His knowledge, passion, care, patience, and valuable academic suggestions helped me to become a mature researcher. I have always been curious and passionate about learning the physics and application of semiconductor materials. Joining Vladimir's group consolidates my mind in pushing the broader of semiconductor material further and investigating more innovative optoelectronic devices. Vladimir's open-mindedness keeps us from developing new ideas and seeking possibilities from those impossible. Plenty of ideas are inspired by him during the routine discussion and presentation between us. His attitude towards life and research encourages me always being better. I really enjoy the time in OneLab due to the endless help from Vladimir.

I would thank my lab mates Mayuran Saravanapavanantham, Jeremiah Mwaura, Benjia Dak Dou, Jun Guan, Richard Swartwout, Karen Yang, Jamie Geng, and Roberto Brenes for their help during our daily experiments and brainstorming sessions. I would like to thank Mayuran especially for his help in training me on all the fabrication and characterization equipment when I first joined the group. His help has built a solid foundation for my later research.

I would like to thank my cooperators: Prof. Moungi Bawendi, and Prof. Brain Anthony, and graduate students/postdocs: Yongli Lu, Shaun Tan, and Meng-Chen Shih in the Bawendi Group. The discussions and cooperation helped me a lot in the study of machine learning-induced next-generation perovskite solar cell devices. I would also like to thank Prof. Sixian You for mentoring me in the experiments in the optics domain. Aside, I'm grateful for the help of Prof.

Sheng Xu at the University of California, San Diego, and my previous group, the Xu group at the Nanoengineering Department. The experience at Xu Group guided me to my life in studying materials and semiconductors.

I would also thank all the staff at MIT. Nano, RLE, and MRL have helped me learn during my research process and trained me in those various equipment and tools.

Last and most importantly, I would like to thank my parents Kai Zhang, Minghui Li, and my wife Yilun Hao for their endless love and support during my life. They make me who I am and encourage me through my life. Especially the accompany of my wife through the time during my undergrad and Ph.D. With her accompany, I have the dare to face any challenges during the time pursuing my dreams and step forward. I dedicate this thesis to them.

Contents

List of the Figures

List of the Tables

Chapter 1 Introduction

1.1 Solar Energy and Solar Cells

Living in a fast-developing world, energy consumption is the foundation of human society and activities. Among all those various types of energy, electricity plays an essential role. According to the research data carried out by "World Energy & Climate Statistics Yearbook 2023", the electrical energy consumption among the whole world has doubled within the last decades. To meet the huge electricity consumption, solar energy, as a clean and renewable energy, has become one of the dominant energies in generating and transferring electricity. According to the statistical data of the U.S. Government, the continental U.S. Total Roof Area Suitable for PV deployment is approximately 4922.3 millions of meter square [11]. Incredibly, the annual generation potential of roof Photovoltaic (PV) installation increases by 25% annually. Thus, if people make full use of the current resources for PV generation, it will be a huge boost for PV-generated electricity.

For the last decades, single junction GaAs solar cell and monocrystalline, polycrystalline, or amorphous Si solar cell has been dominating commercial solar cell fabrication with a maximum Power Conversion Efficiency (PCE) above 45% achieved in the 2010s, also as shown in Fig. 1-1 [12, 13]. However, with the development of solar technology, an increasing number of other photovoltaic technologies have emerged. Materials including optoelectronic materials, organic materials as well as quantum dots are largely investigated in nowadays photovoltaic research. As

Figure 1-1 shows below, starting in the 2000s, the PCE has a steady rise for all those material-based solar cells. Perovskite, with a crystal structure of ABX_3, has a few charming optoelectronic properties that make it a good candidate for next-generation PVs.

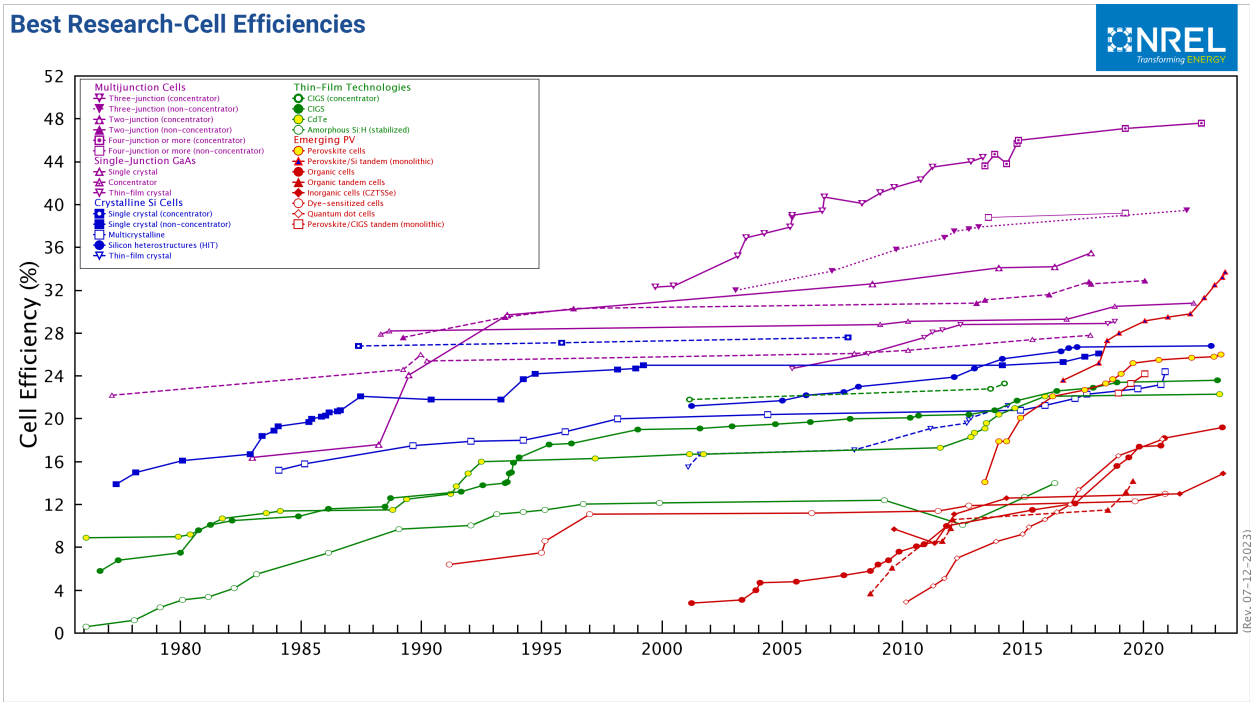


Figure 1-1 Best Solar Cell Efficiency Performance (Source from NREL Official Website)

1.2 Logistics of This Thesis

This thesis focuses on developing the Perovskite PVs towards the goal of commercialization requirements, i.e., Long-term solar cell stability, high PCE, less manufacture cost, as well as scalability.

The work induced in this thesis tries to tackle the challenges in two directions. First, the thesis discusses the interlayer influence of a perovskite solar cell, from the directions of atomic structural information, film morphology and chemistry, interlayer ion migrations, and fabrication limits (Chapter 3). Second, machine learning algorithms are developed to predict the perovskite solar cell photocurrent performance based on the various film optical properties before the device is fully fabricated (Chapter 4). Continuous work will be carried out in the discussion and future work part (Chapter 5).

Chapter 2 Background

2.1 Perovskite Material

2.1.1 Structural Properties

Perovskite is a kind of mineral material that was first discovered in 1839 by mineralogist Gustav Rose and named after Russian mineralogist Lev Perovskite [14]. Perovskite has the atomic structure of ABX_3 with two cations sites A and B with one anion site X. A site is generally considered as organic, inorganic, or metal cations. Some typical metal cations are Rb^+, Cs^+, Fe^+, or organic cation molecules like Methylammonium (MA^+), and Formamidinium (FA^+), tetramethylammonium (TMA^+) cations. B site is typically composed of metallic cations like lead or tin cations with the X site representing common halide group anions like I^-, Br^-, Cl^-, $HCOO^-$, etc. [3, 14-17].

Depending on the various temperature-dependent-phase-change together with the different atomic radii effects, perovskite single crystal or crystallized thin film could form three different structures, as shown in Figure 2-1, cubic, tetragonal, and orthorhombic [5, 14, 18]. Ideally, a cubic structure provides the most symmetric atomic distribution for the inorganic slab. This facilitates the carrier transportation along the Pb-I slab when forming in an optoelectronic device. Changing the radius of the cation or anion is one of the main reasons why perovskite forms into tetragonal forms or orthorhombic forms [5]. In addition to the radii effect, perovskite like $MAPbI_3$, $FAPbI_3$,

or $FAPbBr_3$ is able to form various α phases or β phases depending on the various temperatures during the perovskite synthesizing process under various ambient temperatures [19, 20]. This will also change the lattice parameters which enables various of optoelectronic properties.

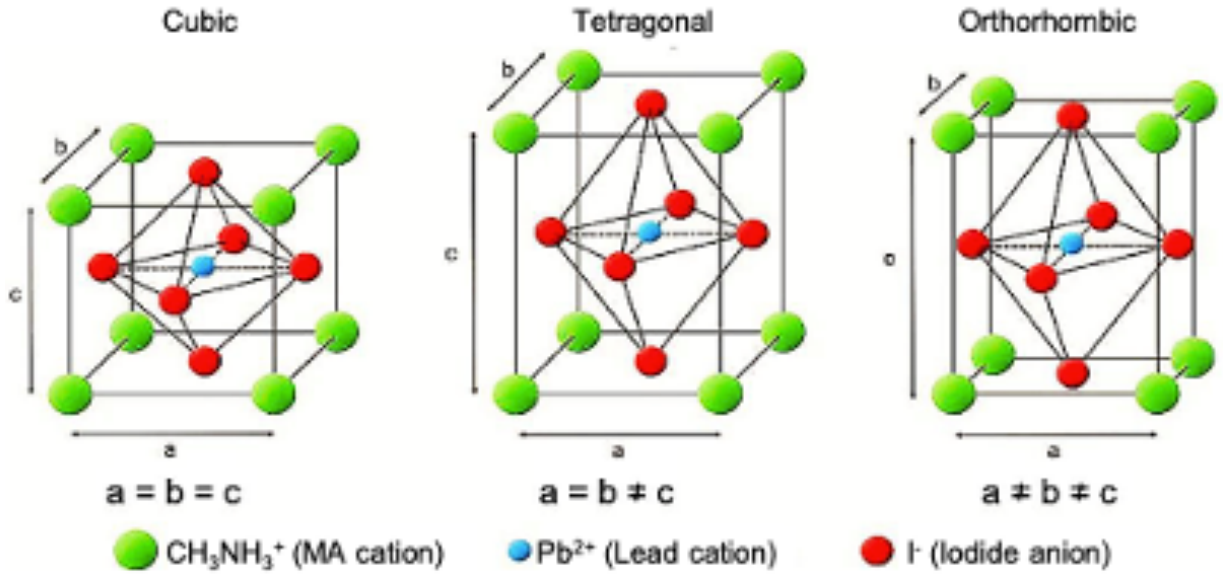


Figure 2-1 General Perovskite structure ($MAPbI_3$ as example) [14]

2.1.2 Optical Properties

Organic metal halide perovskites have shown plenty of extraordinary optoelectronic properties which make them good candidates for various photovoltaic applications [1-4]. The fascinating optoelectronic properties of perovskite largely take credit to their low exciton binding energy, strong light absorption coefficient, relatively long carrier diffusion length, and carrier recombination lifetime [6-8]. All of these contribute to an impressive high power conversion efficiency (PCE) in solar cell devices and make perovskite a good candidate in other optoelectronic

devices like light-emitting-diodes (LEDs), field-effect transistors (FETs) and photodetectors, etc. [7, 21, 22].

Specifically, perovskite has reached a max PCE almost of 26% in 2023 [23]. Due to the various perovskite precursor selections as well as the composition gradients at A, B, or X sites, perovskite can form a tunable electrical bandgap, ranging from 0.87eV beyond 3eV [14]. In addition, by tuning the A site molecule to larger atoms like phenylethyl ammonium (PEA) or butylammonium (BA), quasi-2D Ruddlesden-Popper perovskite is further discovered, forming a multi-quantum well structure with tunable bandgap energy [24, 25]. Aside from the Ruddlesden-Popper structure, quasi-2D perovskites could also form Dion-Jacobson and Aurivillius structures as shown in Fig. 2-2 [26]. In addition, substituting various cations from MA to FA, Cs, Fe, PEA or BA further enlarges or shrinks the perovskite inorganic slab, introducing surface extensions or constraints to the perovskite thin film, which enlarges the property possibilities of perovskite optoelectronic devices [5]. In addition to perovskite's inherent charming optical properties, the solution processability of each layer and deposition accessibility in the corresponding device make perovskite a competitive solar cell candidate against the Si solar cell counterpart [8].

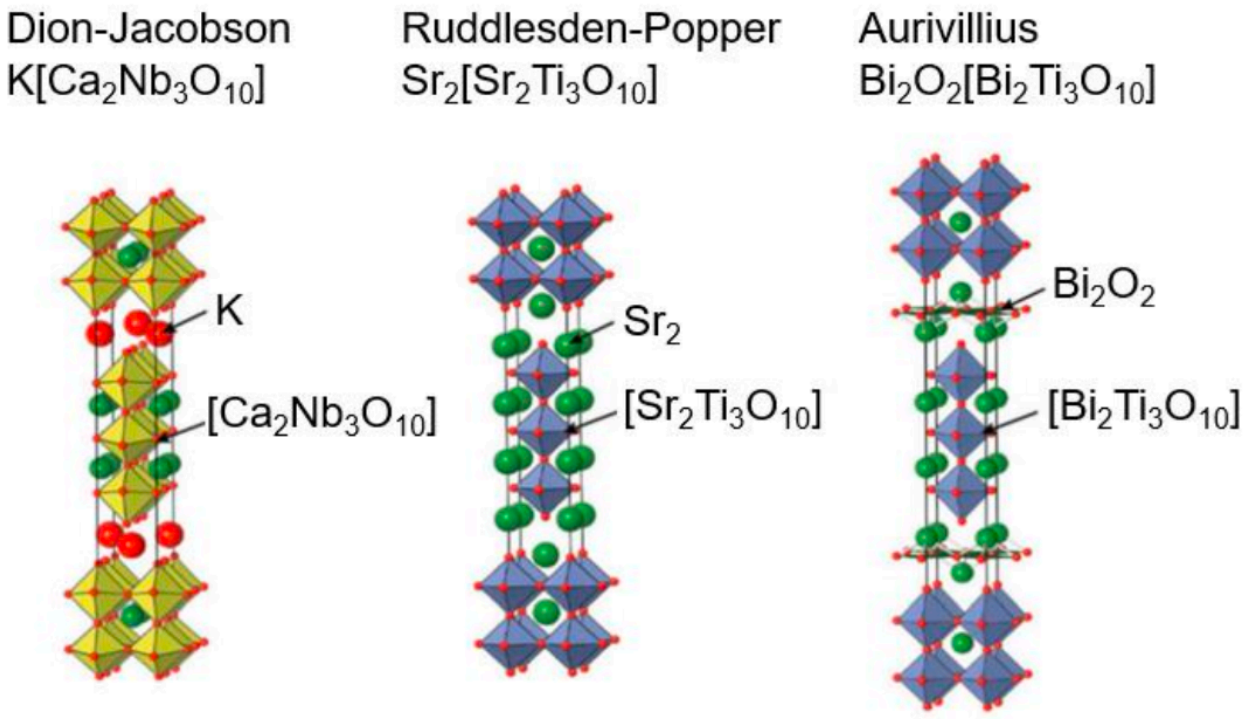


Figure 2-2 Atomic structure for Dion-Jacobson, Ruddlesden-Popper and Aurivillius perovskite. *[26]*

2.1.3 Current Research in Metal/ Organic Halide Perovskite Solar Cells

Metal/ Organic halide perovskite is one of the biggest groups in the perovskite material family. As named, metal halide perovskite consists of an organic/inorganic/metal A-site cation, a metallic B-site cation, and a halide group material C-cite anions. Depending on the level of crystallization, metal halide perovskite is divided into two categories, single-crystalline and polycrystalline. One of the most important characteristics of those two categories is the crystalline morphology difference of the perovskite material.

For single-crystalline, there are no grain boundaries between the perovskite atomic layer. This kind of perovskite is typically synthesized through chemical vapor deposition (CVD), space confinement growth, solution-based crystallization, and solution-based epitaxial growth [14]. An oversaturated perovskite precursor solution is placed under specific temperatures and moistures to form the crystallization nuclei. Depending on the surface energy, the nuclei will continue to grow either in the direction of in-plane ((100) plane) or out-of-plane ((111) plane). Thus, a single-crystalline bulk perovskite crystal is further formed. Additionally, perovskite could also be formed through homoepitaxial growth or heteroepitaxial growth. As indicated by the name, homoepitaxial growth is the growth mechanism by growing the perovskite layer on top of the same perovskite substrate while heteroepitaxial growth has different substrate material and growth material. However, each growth method has its challenges like the difficulties in forming a uniform thin film, crystallization into flakes, and difficulties in controlling the preferential growth sites and thickness. All of these make single crystalline perovskites difficult to assimilate into the requirements of commercialization.

On the other hand, polycrystalline perovskite has plenty of grain boundaries and film vacancies but is relatively easy to fabricate into devices. The most common way to form polycrystalline perovskite is through spin coating continued by thermally annealing. However, polycrystalline perovskite thin films always contain a high density of structural defects. These defects lead to a severe ion migration under biased conditions [27]. In addition, the grain boundaries further limit the carrier mobility inside the films disturb the solar cell performance. In addition, anti-solvents could be used to help recrystallize the polycrystal to form a better thin film morphology [16, 28].

2.2 Challenges towards Perovskite Solar Cell Commercialization

Recently, the reported polycrystalline perovskite-based single junction solar cell has reached a record PCE of above 25% with efficient carrier transportation inside [29]. However, even though reaching such high efficiency, there are still challenges faced by perovskite solar cells to reach the level of commercialization requirements.

First, most of the electrodes still can't get rid of the noble-based electrode (e.g., Au), which cannot prevent degradations of the cell but further enables two-way intralayer diffusion as cell operation time increases. Even though gold has good resistance to corrosion, perfect electrical conductivity, and excellent ductility, the diffusion of gold particles with an out-of-plane direction into the cell enables the degradation of the device: the perovskite layer composed of methylammonium lead iodide ($MAPbI_3$) is likely to degrade into methylammonium iodide and lead iodide under moisture and oxygen [30]. Also, gold particles tend to settle down into the cell which further interacts with the perovskite layer that leaves vacancies behind. In addition to the composition degradation, continuous illumination also degrades the cell, especially with the existence of light in UV wavelength which cannot be filtered out from the incident light source [8, 28, 31, 32]. All these degradation mechanisms prevent a longer cell lifetime under normal environmental conditions.

Secondly, the leakage from the cell is heavily composed of lead-induced compounds that are toxic to the surrounding environment. Due to the defect and vacancies in between each layer,

the intralayer diffusion comes to the source of the ion exchange that enables lead cations to exchange with the air without protection from a denser counter electrode other than gold.

Third, upscaling the cell from an inch-sized sample to a meter-long device tends to trigger more problems including various solution viscosity, perovskite nucleation rate differences, fabrication defects-induced cracks, capsulation difficulties, etc. Additionally, those currently used metals heavily depend on noble metals like gold which largely increases the fabrication cost that could not be utilized for meter-scaled devices. Together, to prevent sputter plasma damage to the deposition substrate (Organic HTL layer like Spiro or long carbon tail polymers), high vacuum thermal evaporation is largely used in laboratory-sized sample deposition of the electrode which is not feasible in large scale. Thus, novel perovskite solar cell structures need to be developed with compatibility of scalable competitive back contact metal electrodes.

Previous researchers have already tried to solve this problem largely in 2015-2018. However, none of them have achieved a solar cell device without dropping its working efficiencies or maintaining a longer carrier lifetime. The trade-off between substituting electrodes with keeping the working properties seems unbalanced. In general, except for the non-scalable carbon-based back contact / HTL (e.g., BCP) layer heterojunctions, there are several substitution materials possible for coating including Silver, Aluminum, Copper, Chromium, Nickle, molybdenum, tungsten, etc. A detailed discussion of each substitution material is presented below.

Silver has been largely discussed as a backup candidate other than gold due to its compatible conductivity and similar work function. However, silver is very active with the halide anions which results in corrosion or contamination of the perovskite layer after a few days of storage. The Ag electrode is reported to react with iodide and bromide anions to form AgI and AgBr molecules through the polycrystalline film grain boundaries and atomic layer vacancies as

diffusion takes place in the direction of outward the cell [31]. Those aggregated AgI particles tend to form a thin layer that acts as an inter-junction barrier that blocks the carrier extractions with huge vacancies left behind, forming block PbI_2 particles which further decrease the conductivity [32]. Thus, even though a roughly same PCE could be reached immediately after fabrication compared with the gold electrode counterpart, that efficiency dropped quickly after days.

Aluminum is also considered as one possible electrode with its abundance on earth. It has a high conductivity with a cheap percent weight cost [28]. However, the oxygen-induced oxidation in air largely hampers Al as a relatively good candidate merely by itself. Various protection layers might be developed to prevent the intralayer diffusion of aluminum cations into the cell towards the perovskite layer. Also, if not encapsulation properly, aluminum oxide might be formed due to the ion exchange from the aluminum layer with the surrounding oxygen particles. Like Aluminum, Copper has a comparable conductivity. However, Copper itself could be quite active interacting with various molecules inside the HTL layer that accelerate degradation. The previously reported stable copper perovskite solar cell requires a C_{60}-induced bi-junction-layered HTL inserted which is not accessible for large-area coating [33]. As for Chromium, it has a high conductivity with a value of $7.9*10^6$ S/m. The addition of Chromium oxide efficiently prevents the degradation from water in the surrounding area. Additionally, the light in weight percent of the chromium layer enables flexible substrate fabrication which facilitates large-area applications towards commercialization. However, the degradation between cathode layer and perovskite still can't be prevented and leads to a band gap work function mismatch further might also contribute to the longtime degradation [34].

Besides the above, Ni is considered a great substitute other than silver and gold electrodes due to its closed working function compared to gold. Additionally, nickel itself is a hard and

condense packed material which inherently decreases the chance of moisture degradation [35]. However, due to its stability and good thermal resistance, it is almost impossible to thermally evaporate onto the solar cell device at a reasonably low temperature which prevents melting the perovskite and HTL layer during deposition. Thus, sputter is largely applied in depositing Ni thin film which is an evitable drawback of sputter plasma damage to the substrate surface. Buffer layers like NiOx might be helpful in decreasing sputter damage but still cannot inherently solve this problem [36]. Owing to these limitations, the efficiency of Ni-based solar cells still can't achieve a level similar to those gold-induced devices. The reported Ni-$MAPbI_3$ perovskite solar cell under n-i-p junction PCE could reach above 10% but is still far from the theoretical Shockley limit [37]. Some other novel III-V metals like Molybdenum is also carried out as possible contact due to their high work function, high electrical conductivity, and chemical stability against halide anions and moisture-induced oxidation. The reported Mo-perovskite device has reached a PCE of 15% which is considerably high compared to other non-noble contact mentioned above [38]. However, the lifetime stability is still undetermined. Also, it is still questionable if Mo could work for single-junction organic-metal halide perovskite devices.

In general, thin film non-noble metal back counter electrodes are still facing the interfacial contacting stability issues, in-plane/out-of-plane pin hole assisted ion migration inwards or outward diffusion of halide-HTL-metal heterojunction layer that minimizes solar cell stability and longtime efficiency. Especially, most of the current approaches seldom take scalability into consideration which is not compatible with our commercialization end goal.

Thus, it is both academia and industrialization interesting to understand the contact working principles and stand inside the loop of scaling to develop a new n-i-p structured perovskite solar cell.

Chapter 3 Perovskite Solar Cell towards Large Scale, Light-weight, and Long-term Stability

3.1 Solar Cell Structure and Band Diagram

A general Perovskite solar cell structure consists of a bottom electrode (typically ITO, FTO, etc.), an electron transporting layer (hole blocking layer, typically SnO_2 nanoparticles), a thin film of perovskite layer, a hole transporting layer (electron blocking layer, typically, spiro-OMeTAD), and a top electrode (typically Au, Ag, etc.). However, with the harsh requirements of gold thin film deposition conditions, i.e., high vacuum-based thermal evaporation, and the non-universally applicable spiro layers in other optoelectronic devices, resulting in a high fabrication cost, it is important to develop a new perovskite solar cell's structure. Based on the foundation of our group's previous work and inspired by other literature [39], we have developed a triple-cation triple-anion perovskite $MA_{0.98}FA_{0.095}Cs0_{.005}Pb(IBrCl)_3$ thin film, with a substitution of spiro layer into PV2001 thin film (Acknowledge to Richard Swartwout's Ph.D. thesis, https://hdl.handle.net/1721.1/140041). Figure 3-1 shows the designed ITO/SnO2+TiO2/$MA_{0.98}FA_{0.095}Cs0_{.005}Pb(IBrCl)_3$ /PV2001/Au perovskite solar cell structures that have been developed in suitable for the large area

coating. To facilitate carrier mobility inside layers, passivation layers might be further needed in between top electrode, p-layer, active layer and/or n-layer.

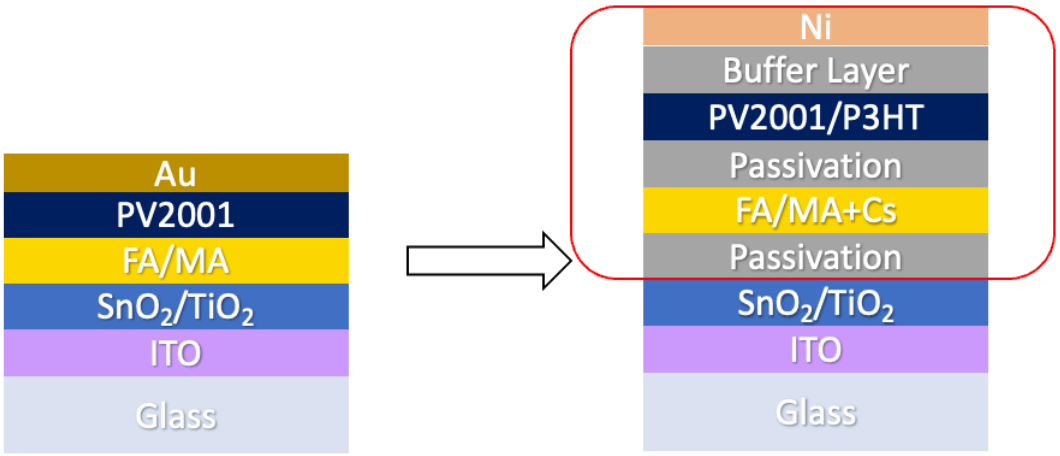


Figure 3-1 Triple-cation/anion perovskite solar cell structure

To further substitute gold top electrodes, various top electrodes w/wo buffer layers are tested. A good Au-substituted contact metal should have four main characteristics: good electrical conductivity, small metal contact resistance, balanced band alignment, and operational stability. In general, except for gold, silver, copper, aluminum, zinc, nickel, and magnesium have relatively good electrical conductivity. In addition, Ag, Cu, Al, and Ni have shown a match work function (around 4.2 to 4.6 eV) with a low contact resistance [40].

Metal	Al	Ag	Au	Ni
Contact Resistance (Ω)	30.31	0.96	0.87	1.27
Contact Resistivity ($m\Omega*cm^2$)	2051.71	35.64	34.28	46.91
Work Function (eV)	4.28	4.26	5.1	5.15

Table 1 Metal Contact resistance, resistivity, and work function

Figure 3-2-a shows the band diagram of the solar cell structure. Effective carrier transportation should allow negative charges to be extracted from the perovskite active layer into the n-layer further collected by the electrode with the p-layer blocking the charge to make a one-way charge drifting tunnel in the lowest unoccupied molecular orbital (LUMO) at the band edge. A similar design but with opposite energy level is followed for the positive charge located in the highest occupied molecular orbital (HOMO) at the band edge (Figure 3-2-b).

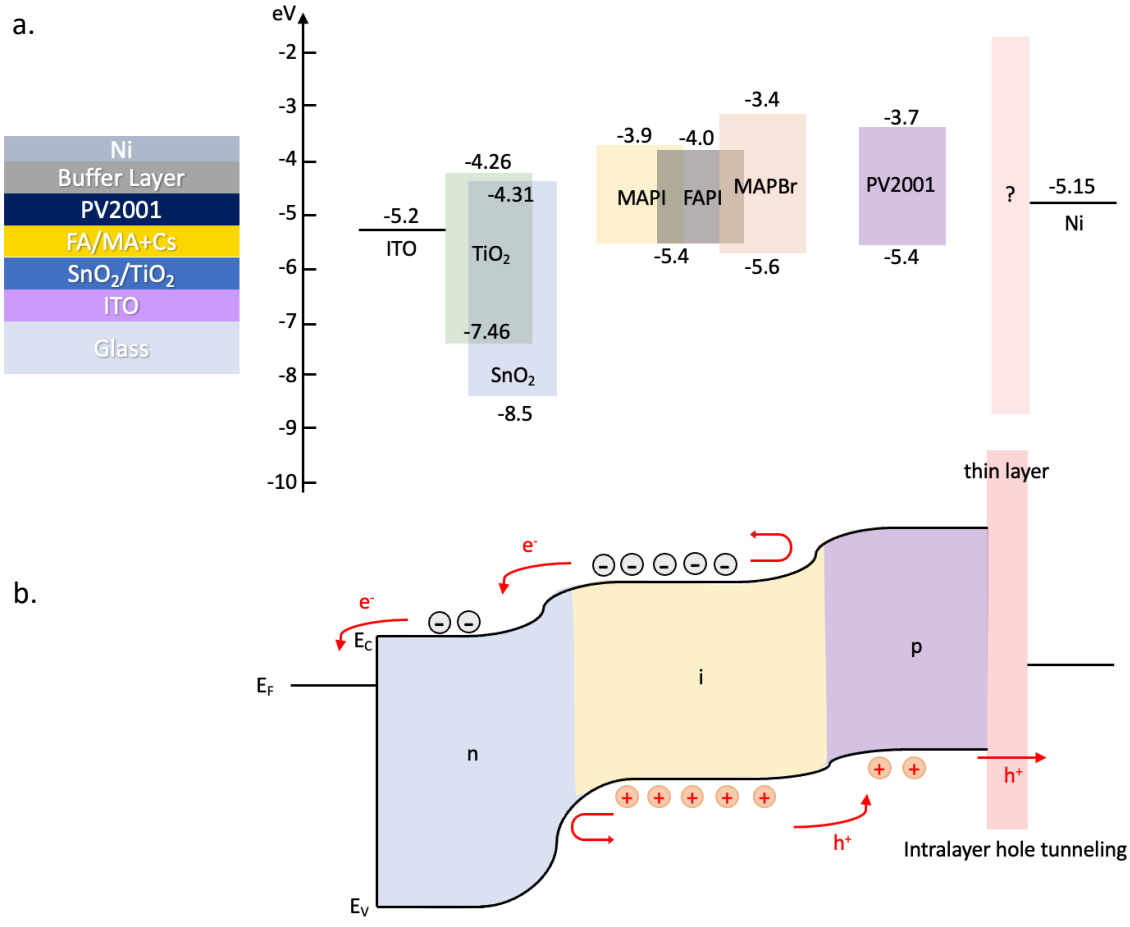


Figure 3-2 a. Perovskite solar cell structure bandgap energy with electron-hole pairs intralayer diffusions and drifts. b. Effective intralayer charge transportations in a typical n-i-p diode.

Ideally, a buffer layer is despised as it prohibits the intralayer carrier transportation by either introducing more vacancies or traps inside the thin film. It also introduces one more interface (buffer layer-metal interface) that could increase the roughness of the device. However, substituting gold with other heavy metals in a mild environment requires a switch of the deposition method from high vacuum thermal evaporation into a more industrial-matured sputter system. A target metal substrate is hit by the ionized argon or nitrogen gas inside the sputter chamber under a relatively medium vacuum level (10^{-3} Torr instead of 10^{-6} Torr for thermal evaporator). In the meantime, the sputtered atom will be released from the source target and deposited into the sample substrate. Along with the whole process, ionized argon molecules release those extra negatively charged oxygen atoms with the addition of electrons that lead to intense UV damage to the sample substrate. In addition, the energy carried from ionized inert gas molecules could further cause sputter damage that bumps into the sample substrate and breaks the atomic bonding on top of the sample substrates. Thus, to protect the underneath structures from sputter plasma damage, it is important to have a buffer layer.

Need to note that the thickness of the buffer layer plays an essential role in designing the new solar cell structure. To facilitate carrier mobility inside the diode, electron-hole-pair tunneling could be achieved only when the buffer layer is thin enough. A thick layer increases the shunt resistance and the series resistance that decrease the short-circuit current, leading to decrease in solar cell's PCE. Thus, the buffer layer's thickness will also be discussed in the later part of this thesis.

3.2 Material Selection and Thin Film Design

The new design could be generally divided into two parts, directly substitute gold with other metal or with the help of a buffer layer. The buffer layer could be generally divided into three categories, metal oxide, metal fluoride, and organic large molecules. Some general metal oxide used as buffer layer includes Al_2O_3, MoO_3, ZnO, etc., and metal fluoride includes LiF, MgF, etc. For the organic large molecules, utilizing those most common organic Light Emitting Diodes (OLEDs) materials, Figure 3-3 shows a collection of good holes transporting large molecule polymer materials with the aligned bandgap following the effective carrier transportation.

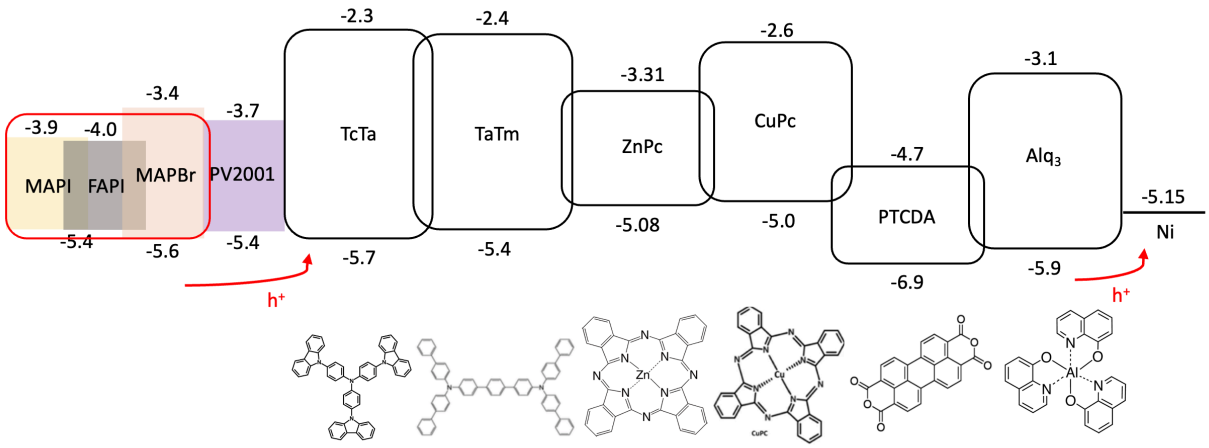


Figure 3-3 Possible Intralayer buffer materials with the corresponded bandgap energy.

3.3 Heavy Metal-based Perovskite Solar Cell Contact

Silver as one of the most common substitutes for gold in optoelectronic applications is studied first. A controlled baseline solar cell structure is determined as ITO/SnO_2 + TiO_2/$MA_{0.98}FA_{0.095}Cs_{0.005}$ $Pb(IBrCl)_3$/PV2001. A commercially purchased one-inch ITO-glass substrate is cleaned and UVO-zoned followed by spin coating the rest three layers and thermally annealed on the heat plate. A current-voltage sweep is used to determine the properties of the as-fabricated solar cell shown in Figure 3-4. Figure 3-4a shows the controlled group applying gold as back contact. A similar JV property is discovered by merely substituting gold with silver by observing the open-circuit voltage (Voc), short-circuit current (Jsc), Fill Factor (FF), and final Percent Conversion Efficiency (PCE) (Figure 3-4 a, b). However, a low cell stability is discovered for the silver electrode after storing the silver for 3 days and 7 days (Figure 3-4 c, d). This is because of the fundamental defects of polycrystalline perovskite thin films. The grain boundaries in the polycrystalline perovskite layer facilitate the out-of-plane ion migrations that extracts the atom from the active layer, reacting with the top silver electrode that further decay the cell. A SEM image of the top view of spin-coated perovskite thin film is shown in Figure 3-5. Particle clusters are clearly shown in the pictures, with the grain boundaries largely distributed all around the perovskite thin films.

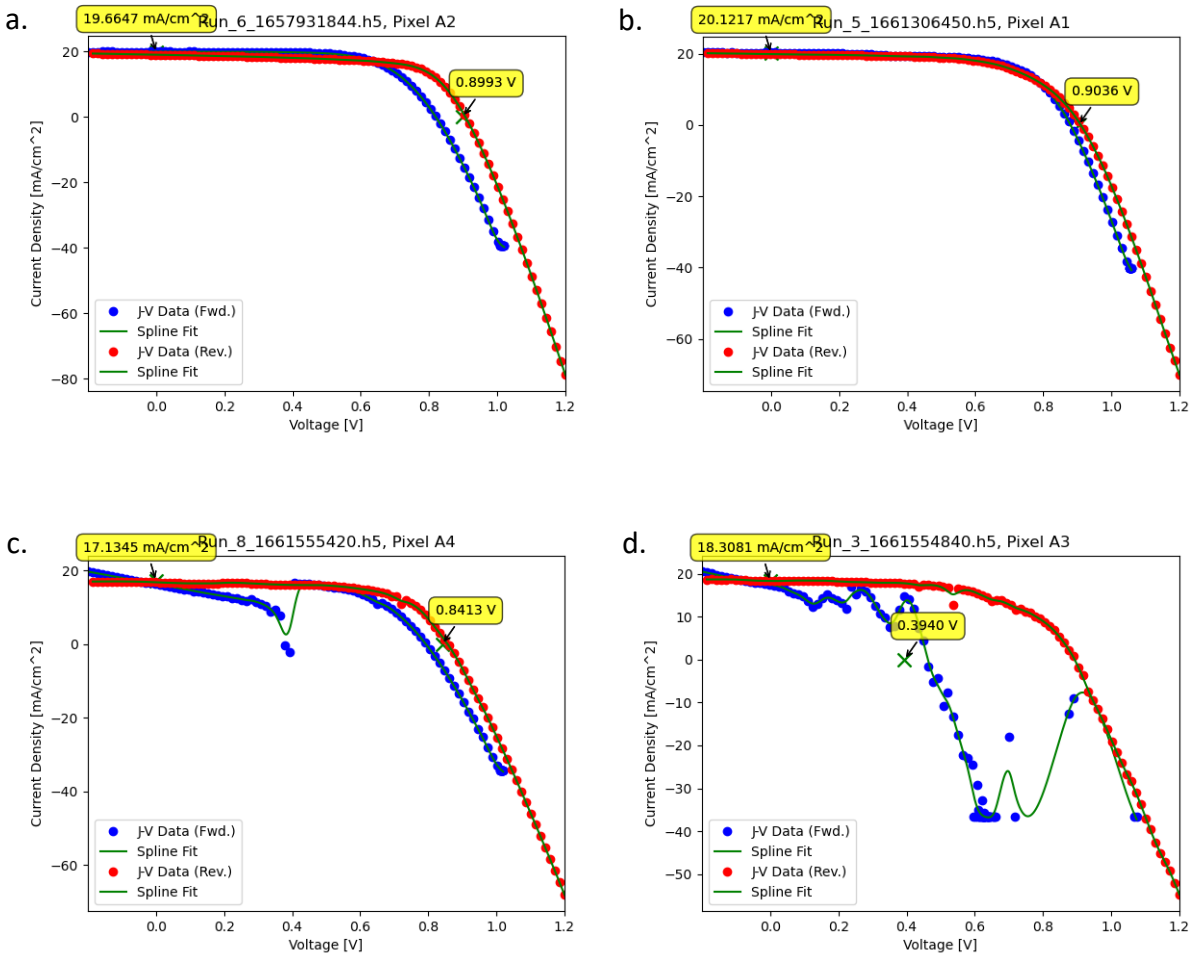


Figure 3-4 Perovskite JV-sweep based on a. Au electrode, b. Ag electrode, c. decayed Ag electrode after 3 days, d. decayed Ag electrode after 7 days.

Especially, due to the grain boundaries and vacancies on top of the bare perovskite film, while evaporating the Ag particles, the silver particles could form multiple clusters on top of the rough top surface grains and start to sink through the boundaries further penetrating all the way through the PV2001 layers, forming AgI or AgBr that exchanged anions from the perovskite layer. Meanwhile, HI and I_2 will be released into the surroundings through the vacancies that further damage the cell surface. Those formed Ag-compounds will residue on top of the solar cell leading to a drop in FF and increasing the series resistance.

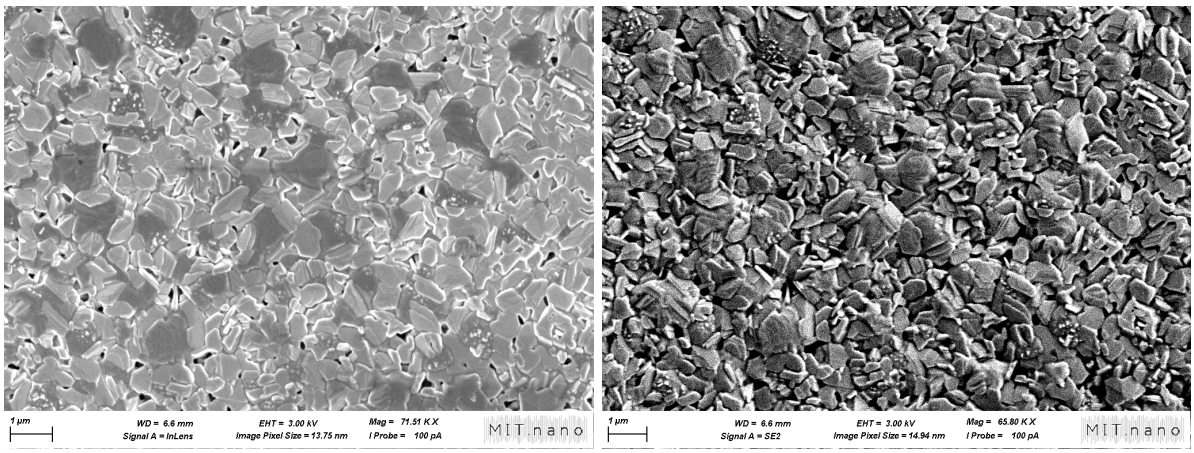


Figure 3-5 SEM image of polycrystalline perovskite thin film

A theta-2 theta XRD scan is carried out to further determine the silver-based solar cell degradation process. Characterization peaks corresponding to silver (111), (200), and (220) are clearly shown at around 38º, 44º, and 64º degrees in the fresh cell. However, after 7 days of degradation, the aged cell shows a peak with a shift towards smaller diffraction angle due to the lattice expansion with strain relaxation in the crystal. Additionally, the peak intensity decreases due to the perovskite crystallinity decreases. There might be single crystalline perovskite formed due to the long-time storage. Notably, except for the silver peaks, side peaks turn to show up representing (111) AgI, (220) AgBr, as well as (331) AgI at 24º, 40º, and 60º on top of the perovskite surface. This further demonstrates the degradation of Ag in reacting with the halide group in perovskite layer. Thus, even with a good optoelectronic property in substitution of silver, it is still problematic due to the solar cell's long-term stability.

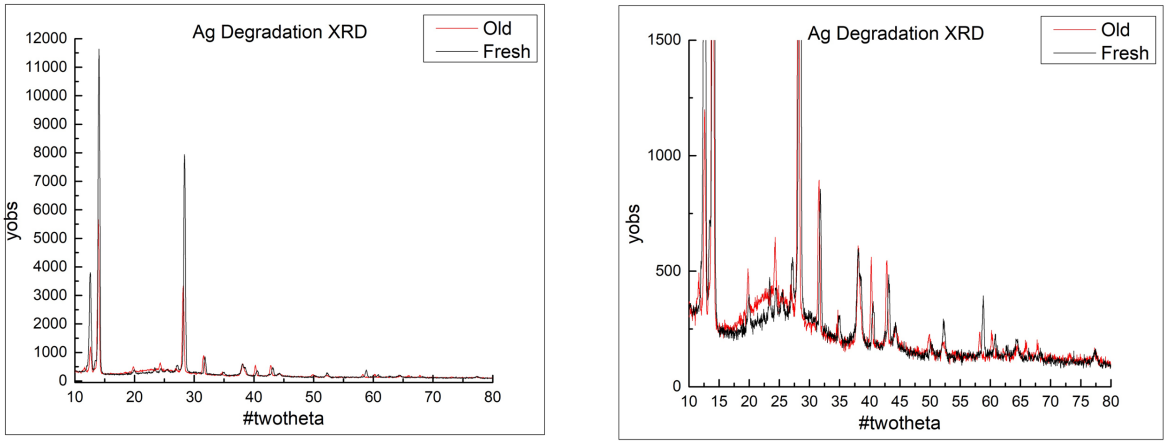


Figure 3-6 XRD Pattern for aged and fresh perovskite solar cell based on silver top electrode (with zoomed-in version).

Except for silver, single layers inducing other metals like Al, Ni, and Cu happen not to be good candidates in substituting the gold either due to the high reactivity of the metal that reacts with the substrate materials or through sputter plasma damage. The JV-sweep corresponding to these designs is presented in the appendix section below and won't be largely discussed here.

3.4 Buffer layer-induced Perovskite Solar Cell

Combinations like LiF + Ag/Al/Cu/Au, MoO_3 + Ni/Al, TcTa + Ni/Au, and ZnPc + Au are further investigated in this section. After systematic research, sputtered or E-beam evaporated Ni, Al, and Cu electrodes tend to have a non-diode curve effect due to the sputtering damage and energy induced by the electron beam. As shown in Figure 3-7 below, the top view SEM image shows that sputtered nickel film tends to have a more uniform surface morphology with fewer grain

boundaries compared to the thermally evaporated silver electrode. Together, as Ni is more chemically stable and mechanically tough than Ag, nickel is selected as the top counter electrode in the new structure design.

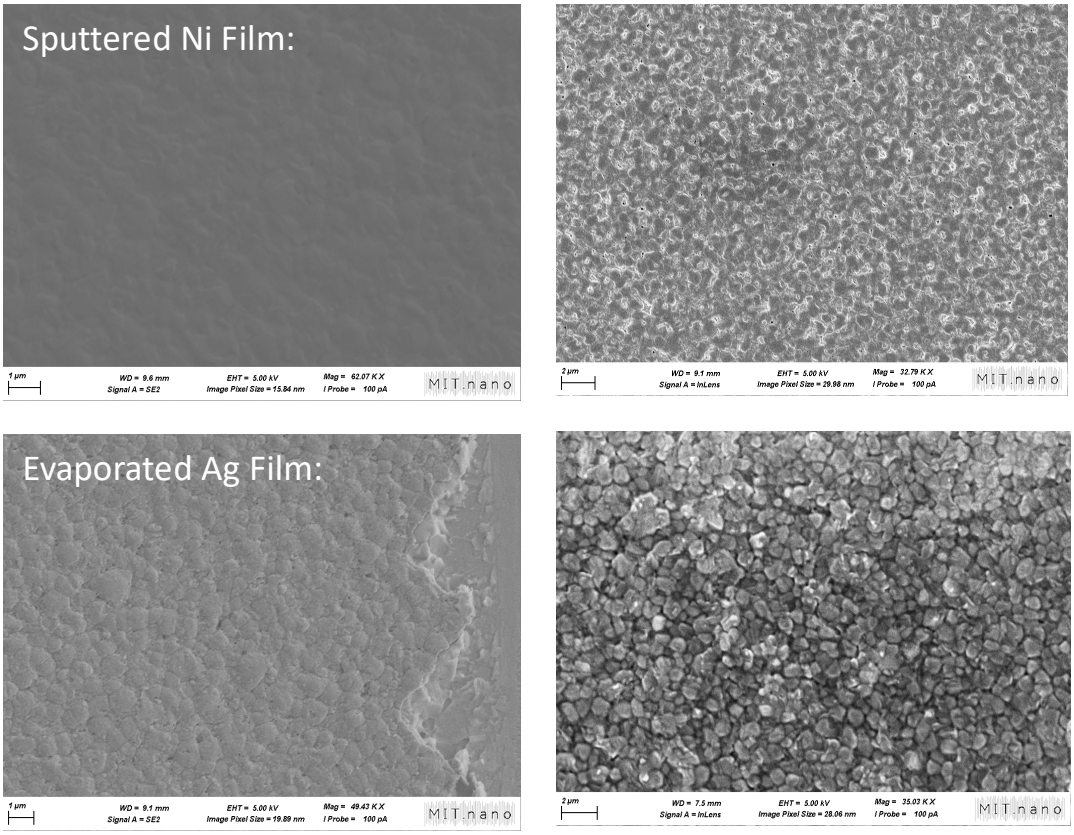


Figure 3-7 Top SEM image of sputtered Ni thin film and thermally evaporated Ag thin film.

Lithium fluoride and molybdenum oxide are developed as representatives of metal oxide and metal fluoride to solve the deposition challenges. For LiF-based structure, various metals are again tested as contact. Aluminum deposition is unstable in both thermal evaporation and sputter due to the high reactivity of metal leading to a fast thermal heat-up and cool down process. Copper is reactive as it approaches the solar cell substrate, it melts the substrate, especially the thermally

sensitive perovskite active layer. Despite the variation of metal contract, all solar cells including lithium fluoride turn out to have a low Voc, Jsc, and FF due to the LiF's low conductivity. For the MoO_3-based buffer layer, the as-deposited solar cell tends to have an acceptable but lower Jsc with a low FF. This is due to the high valance band level (9.6 eV) which forms a Schlocky barrier layer that reduces the electron-hole pair intralayer carrier extractions. A set of detailed graphs are shown in the appendix.

The organic large molecule buffer layer is further studied to have better carrier mobility. Tris(4-carbazoyl-9-ylphenyl)amine (TcTa), with a matching HUMO and LOMO of 2.3eV and 5.7eV, turns out to be a good electron blocking layer and hole transporting layer resulting in a comparably best photovoltaic performance out of other buffer components. With a combination of 3 nm TcTa in contact with PV2001 and 70 nm Au contact, a PCE of 13.42% is reached. As a comparison, the best cell fabricated based on Au contact under this commercialization structure is 15.62%. This is shown in Figure 3-8a below. An extended experiment shows that TcTa is able to protect the substrate layers from sputter plasma damage during the sputtering process. Under the same baseline structure, gold is substituted for Ni as a back contact, and a 3.35 percent PCE drop is discovered (Figure 3-8b). A stability test is also carried out in Figure 3-8c as the solar cell is stored under nitrogen and room temperature in simulating the encapsulation situation. After two weeks of storage, FF tends to decrease the most while slightly decreasing the Jsc and Voc, leading to a decrease in overall PCE of 5%. Figure 3-8d shows the final PCE comparison of the JV sweep under forward bias after each cell is freshly made. Need to note that for the freshly made cell, hysteresis is slightly enlarged compared to the pure Au or TcTa/Au counterparts as the buffer layer's thickness is not optimized that ion migration could be slightly blocked (as the thickness of TcTa layer increases from 3 nm to 5 nm, which enlarges the series resistance and weakening the

hole tunneling transportation through the buffer layer). Hence, to further prove the influence of the TcTa thickness on the perovskite's solar cell photovoltaic properties, thickness-controlled experiments are carried out.

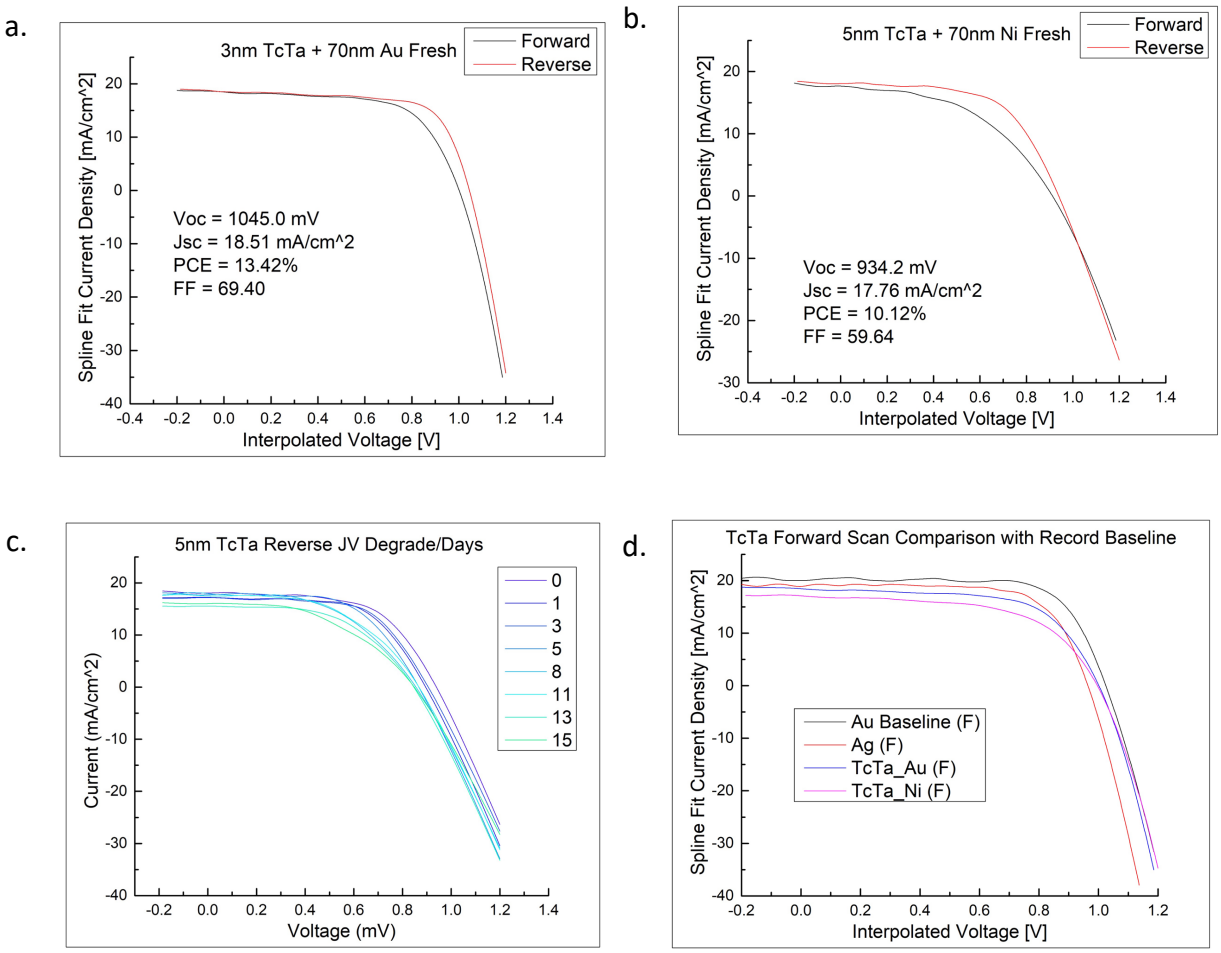


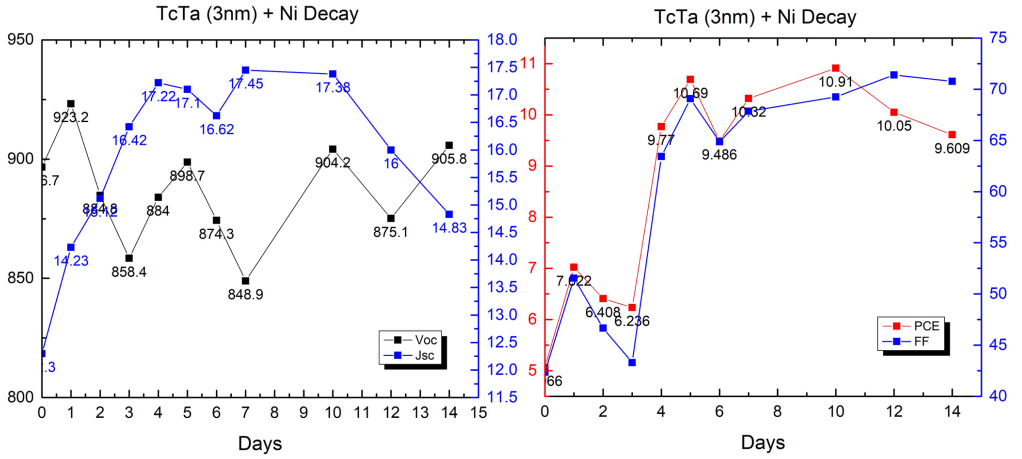
Figure 3-8 JV-Curve of a. 5nm TcTa + 70nm Au Perovskite Solar Cell, b. 5nm TcTa +70nm sputtered Ni, c. TcTa + Ni solar cell degradation study, d. Forward scan JV-curve comparison under different circumstances.

Figure 3-9 shows the Voc, Jsc, FF, and PCE comparison based on 3 nm, 5 nm, 8 nm, and 10 nm thick TcTa buffer layers. It is found that the thinner the TcTa films, the better the overall

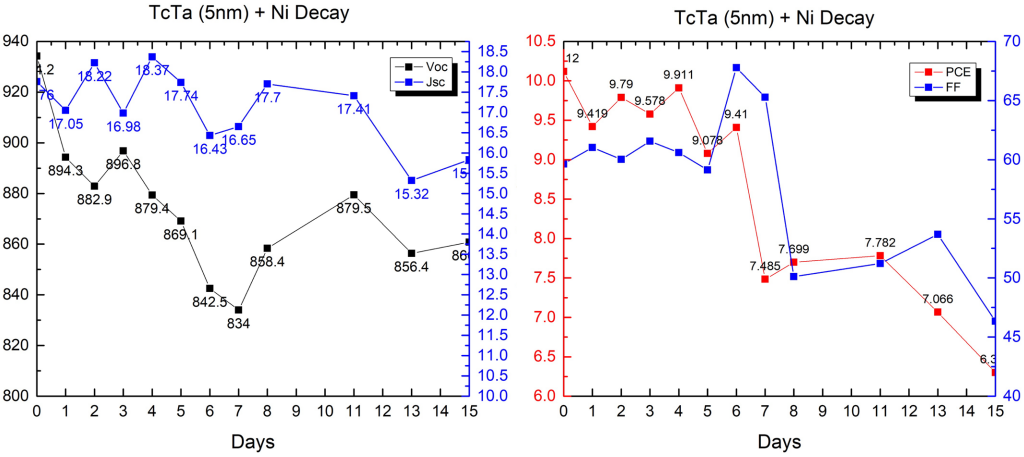
solar cell performance could reach. However, 3 nm is found to be a threshold that further below 3 nm could not protect the substrate structures from the plasma sputter damage. Another edge case is also studied that without the layer of PV2001, 100nm TcTa is thermally evaporated both as the HTL material and as the buffer layer. However, it turns out to have a Voc of 861mV, Jsc of 16.58mA per square centimeter but a low FF with only 45% (shown in the appendix). Other materials like ZnPc, and Alq_3 are also tested but all with a lower performance compared to TcTa.

Notably, for TcTa in contact with gold, after stabliizing the cell for 3 to 4 days, a boost in Jsc is discovered with the FF and PCE all increasing. One possible reason is due to the deposition influence that immediately after deposit, it tends to form a “hot surface” for those thermally evaporated buffer layers and top electrodes that inhibit the ion transportation. Another possible cause is due to the amorphous structure of the as deposited surface that tends to have larger vacancies or grain boundaries immediately after deposition. After times, all the coated surfaces are cooled/settled down and could be more stabilized with those microporous vacancies further fulfilled forming a more uniform surface morphology due to the influence of gravity or the inert surrounding environments.

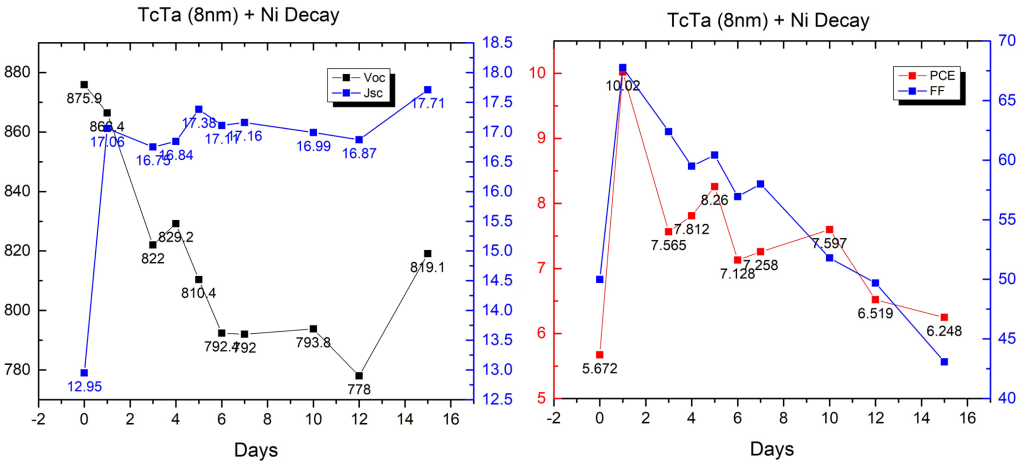
a.


TcTa (3nm) + Ni Decay
Voc
Jsc
Days
TcTa (3nm) + Ni Decay
PCE
FF
Days

b.


TcTa (5nm) + Ni Decay
Voc
Jsc
Days
TcTa (5nm) + Ni Decay
PCE
FF
Days

c.


TcTa (8nm) + Ni Decay
Voc
Jsc
Days
TcTa (8nm) + Ni Decay
PCE
FF
Days

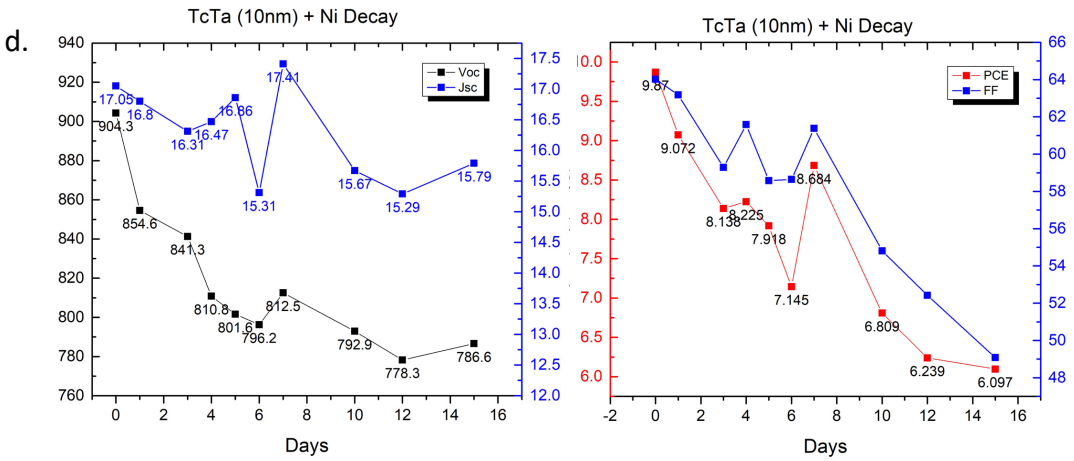


Figure 3-9 TcTa Thickness Comparison for a. 3nm, b. 5nm, c. 8nm, d. 10nm.

Additionally, to further compare the difference between 5 nm TcTa and 3 nm TcTa, percent quantum efficiency is also measured and plotted together with the PCE degradation curve shown in Figure 3-10. As shown in the EQE curve, the %QE between 400 nm wavelength to 800 nm (broadband %QE) tends to increase when substituting 5 nm TcTa with 3 nm TcTa. A maximum %QE of 93% is discovered corresponding to the wavelength range of perovskite materials.

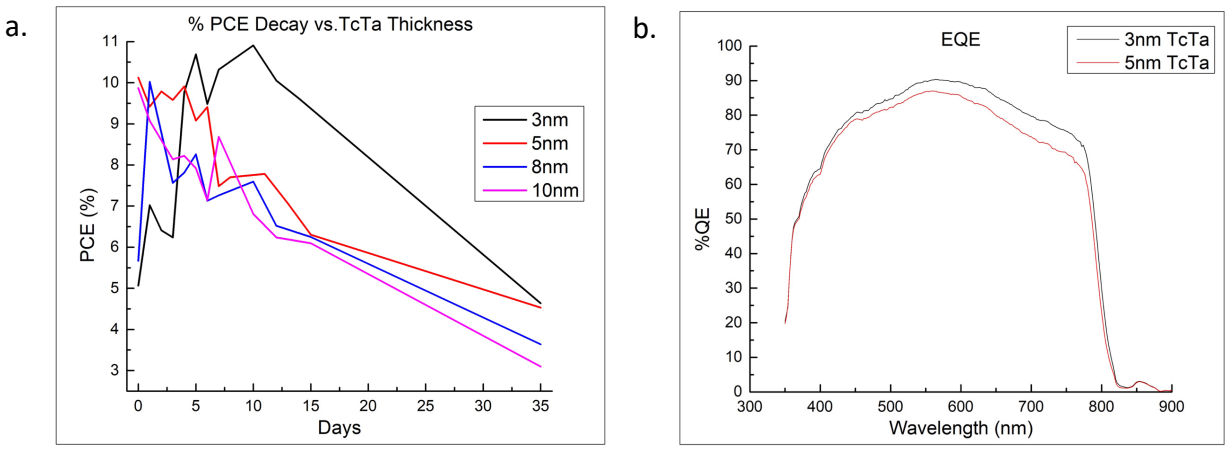


Figure 3-10 a. PCE Decay vs. various TcTa thickness. b, Percent QE comparision based on 3 nm and 5 nm TcTa.

In all, TcTa is found to be the most promising material as a buffer layer to prevent sputter damage. 3 nm is the best buffer layer thickness from the point of view of both enabling intralayer carrier tunneling and proper band alignment. 3 nm TcTa tends to have the best photovoltaic properties like PCE, and long-term stability compared to the other thickness or buffer layer counterparts. A comparison to other structures' solar cell counterparts is again shown in Figure 3-11. With the additional layer of 3 nm TcTa, less than 1% drop of PCE is discovered which further enables the gold-free structure.

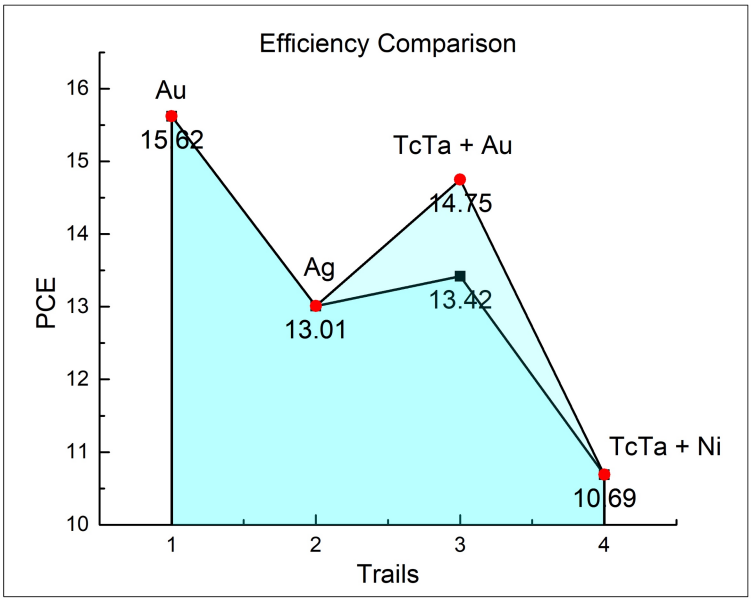


Figure 3-11 Comparison of TcTa-based system with other structure baseline perovskite solar cells.

Chapter 4 Perovskite Solar Cell Devices Optical Performance Prediction through Machine Learning

4.1 Machine Learning in Perovskite Optoelectronics

Organic-metal halide perovskites have promising optoelectronic properties, making them outstand in next-generation photovoltaic devices. However, the performance of solar cells largely depends on the fabrication process or various film recipes [41]. With the fast-developing world, artificial intelligence shows up as a powerful tool to help researchers to analyze, train, simulate, and further predict the experimental data, which largely save the research time and cost in those physical-heavy-based projects [41, 42]. Starting at 2018, an increasing number of perovskite solar cell research starts to have machine learning in their publications. The portion number of machine learning-induced perovskite publication has further increased to almost 50% of all perovskite's optoelectronic application publications in the year of 2022 [42, 43]. In general, data collection, data processing, model development, model evaluation and model application are the five fundamental steps in machine learning algorithm [44]. The required number of samples needed to train reasonable ML models is typically more than three times the controlled parameters, with errors in data acquisition not taken into considerations [44]. Utilizing the advantages of processing

a great deal of data, researchers have applied machine learning to investigate thermal stability [45], select capping layers to suppress perovskite degradation [46], discover integrated-back-contact material thin films [47], and to predict optical behavior [43]. However, even with those dedications in promoting machine-learning induced perovskite solar cell studies, it is still challenging in predicting the solar cell's photovoltaic performance with a high regression accuracy [48].

Under the roadmap of transferring laboratory perovskite solar cell fabrication method into industrialization method, challenges like ink selection, stability enhancement, labor, or time cost, etc. are inevitable. Specifically, most of the laboratory-sized record cells are not applicable in transferring to a printable method like slot-die coating or roll-to-roll coating, etc. Hence, it is essential to predict the possible performance before head while investigating various novel perovskite inks [47].

4.2 Machine Learning First Work and Limitations

Thanks to the first work in our lab exploring physics-based models using the work detailed by Brenes et al. in Phys. Rev. Applied 12, 014017 (2019) [49], the basis is established through a model that accounts for losses due to bulk Shockley-Raidhall (SRH) and Auger recombination, and gains due to photon recycling. This model helps to assess the maximum potential power-conversion efficiency (PCE) for perovskite thin films. Figure 4-1 shows the fits to intensity-dependent TRPL data collected on $FAPb(I_xBr_{1-x})_3$ films passivated with hexylammonium bromide. Through the recombination rate constants extracted from these fits, these passivated films are

determined to have a maximum potential PCE ~28% while the devices have achieved maximum efficiencies of ~25%, indicating only small amounts of loss due to the presence of either transport losses during charge extraction or interfacial recombination.

However, the model presents limitations for predicting actual devices, as it does not incorporate the interfacial recombination between layers in the whole device stack, band-bending and internal electric fields developed due to the charge extraction layers and charge transport losses. These non-idealities could be included using a drift-diffusion model and using fits to experimental data to determine the relevant simulation parameters. Nevertheless, the number of free parameters (degree of freedom) in these models require extensive measurements and fitting models that hinders their usefulness as a predictive tool.

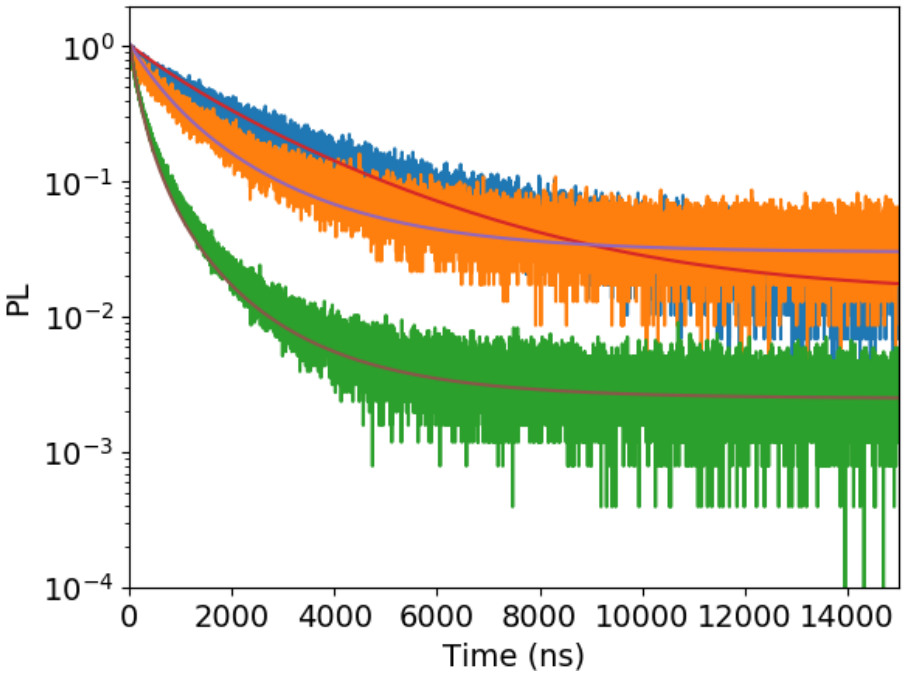


Figure 4-1 Intensity dependent TRPL data taken at carrier densities 1014 cm-3 (blue), 1015 cm-3 (orange) and 1016 cm-3 (green). The solid lines represent fits to the data using a rate equation considering SRH, radiative and Auger recombination.

Therefore, I have begun exploring the use of machine learning to take the measured physical parameters of the cell as input and directly predict the JV curve, thereby predicting the cell efficiency performance. These Blackbox models allow us to potentially predict the performance of full devices with data that might contain many physical processes that could be too complex to predict accurately with a physics-based model. Hence, we consider using Time-resolved photoluminescence (TRPL), Spectrum-resolved photoluminescence (SRPL) as well as percent transmission (%T) curves as inputs to investigate the relationship underneath. A logistical step of the whole project is presented in Figure 4-2 below.

Figure 4-2 Machine learning-induced prediction model development steps.

4.3 Device Data Acquisition

Thanks to the help of Optigon, Inc, we could obtain all three curves (SrPL, TrPL, % Transmission) within 3 seconds per cell. The sample collection process is shown in Figure 4-3. This commercial

set-up is based on the lab-version set-up from Brandon Motes's development with a faster acquisition rate and multiple-sample acquisition capacities (https://hdl.handle.net/1721.1/147276). Samples are located on the sample holder with a moving rod on both sides. Two different detectors are placed on top of the system, acquiring optical signals by sequence.

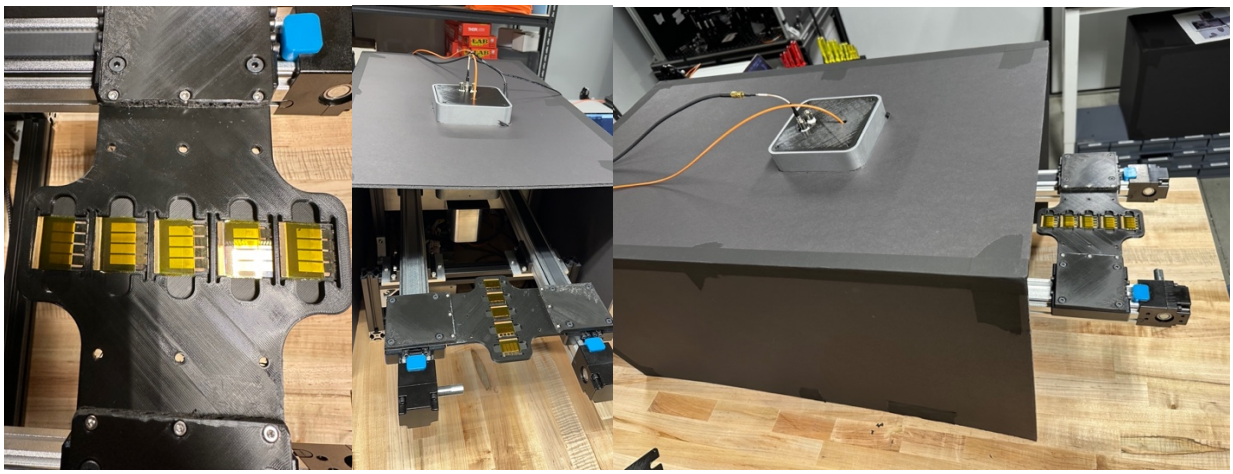


Figure 4-3 Percent Transmission, SRPL and TRPL Data collection set-up by Optigon, Inc. Five samples are collected as a batch located on a moving trail with a total collection time less than 20 seconds in total.

4.4 Neural Network Regression Algorithm

To start with the prediction model development, a few important assumptions need to be declared. The first assumption is that all four cells in a single batch are consistent enough that the JV curve measured from a completed device in a batch serves as an accurate ground truth for a stack that was prematurely removed from the fabrication process, leaving it with only the layers up to and

including the perovskite. This assumption allows us to generate data from current device fabrication procedures with minimal deviation.

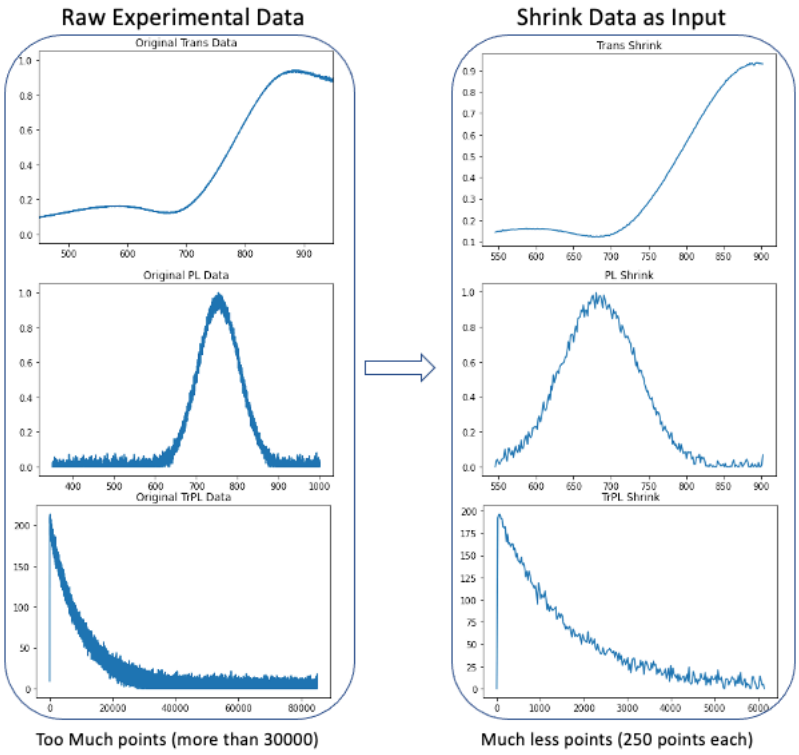


Figure 4-4 Shrink of input Degree of Freedom (DOF) of the Learning Model.

The second assumption is that a diode circuit equivalent model is a suitable model to build the architecture around. As the input TRPL, SRPL and Transmission data are 2*N arrays, where N is the number of bins in the histogram, the input is effectively a picture. Likewise, the output JV curve is an N*3 array. This gives an extremely high number of degrees of freedom (DOF) for the model. Typically, each measurement takes more than 3000 points with even more than 10000 points obtained through TRPL. This is not appropriate for any model because this subtask's largest constraint is the amount of data available for training. Initially, we plan to test using the raw data

directly instead of fitting for decay coefficients but will still consider fitting, which would leave us with an extremely simple neural net going between a few scalars as inputs and outputs. It was our intuition that fitting the trace first would reduce the benefit of allowing arbitrary physics to be represented that is provided by using ML and would force our system to be bound to expected physics. Hence, as an eclectic method to reduce the DOF, all the data are constrained within the wavelength range of 550-900 nm which perfectly aligns most of the perovskite solar cells. Besides, as the raw data curve has a small acquisition interval, we choose to reduce the data point density by extracting one point per three neighborhoods (the one which divide 3 mode 1). Hence, after processing the raw data, we have three input matrices with 250 data points each related to %T, SRPL, and TRPL correspondingly (Figure 4-4).

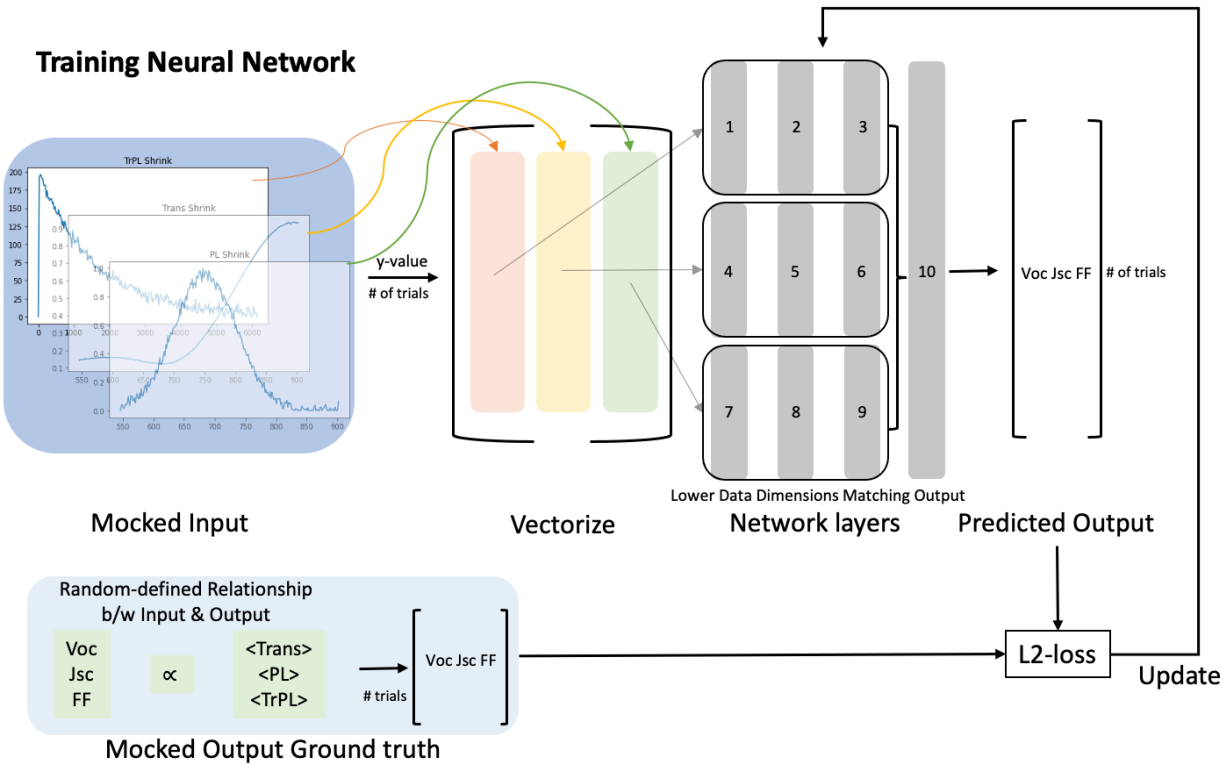


Figure 4-5 Architecture of Neural Network based Machine-learning Prediction.

Among all those various Machine-learning model published, neutral network is good at seeking possible relationships inside the Blackbox between the physical property inputs and the device performance output (here, value of open-circuit voltage, Voc, short-circuit current, Jsc and fill factor, FF). A neural network architecture is developed as shown in Figure 4-5. The three 2*250 matrices representing three input graphs are concatenated into one 2*750 matrix. To first prove the learning model is able to learn from the input and could pass the validation dataset, we have hardcoded plenty of mock data. After vectorizing all the sample into a three-dimensional matrix, each submatrix sized m*2*250 represents all 250 datapoints from %T, SRPL and TRPL. Here, m represents how many samples we put into the algorithm. Each submatrix goes through a three-layer-model-training which downgrades the learning dimension from 250 to 100 to 50 further to 3. Within this forward process, a three-dimensional predicted value is then generated from each submatrix input. Another concatenated layer afterwards helps to downgrade again the 9 inputs (3 from each curve) further to our 1*3 vector output representing Voc, Jsc and FF value. During the training process, the m groups of Voc, Jsc and FF value is imported to the model as the ground truth to minimize the L-1 or L-2 norm in the lost function under reverse process.

Need to note that three important parameters are investigated to facilitate a better training output ---- number of input data, gradient descent steps and number of epochs. Overtraining the model with excess number of epochs or input data may result in overfitting the model while undertraining the model may achieve a less accurate prediction result. Hence, it is important to determine the available epoch range for training model to be converged. For a specific dataset, number of epochs and gradient descent steps could vary a lot. Possible ranges are 1000 to 100000 and 0.000001 to 0.1 correspondingly. Need to mention that the more epochs, the longer time it

takes to train the model. To further investigate the influence of number of input samples, the mocked data are generated in a group of 1000, 100 and 30. Under all three groups, the algorithm could learn perfectly under L-1 or L-2 norm reverse lost function. And the final converged stepwise error all appears a decrease for 10^6 times (Figure 4-6) reaching a L2-norm value of 0.3-0.5. This convinced that even with small perturbation, as the dataset is appropriate to the model, i.e., the characteristics of the input are sufficient for the model learning, number of inputs won't have a huge influence on the algorithm.

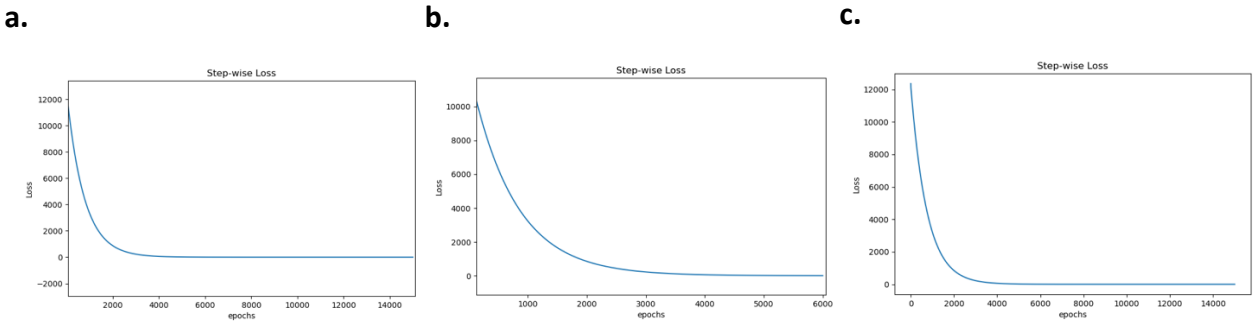


Figure 4-6 Stepwise training loss based on a. 1000, b. 100, c. 30 number of mock-sample inputs. The converged loss is 0.43, 0.41 and 0.30 correspondingly.

To start model learning, we develop three groups of data: one group of mocked data containing 100 training samples and 10 validations samples; two groups of real data with one group containing 20 training data with 4 validation data and the other with 212 training data with 8 validation data. Starting with the two smaller dataset, validation results matrices are generated. Figure 4-7 shows the stepwise training and validation loss based on the 20-sample dataset. This proves as an example that the model converges perfectly to a relatively low error range. Prediction results are generated and shown in Figure 4-8 below based on the 20 experimentally acquired samples.

Figure 4-8a is the measured ground truth output for the validation dataset which is used as a reference to calculate the percentage error. After training the model for 20000 epochs, predicted Voc, Jsc and FF value is shown in Figure 4-8b. Here, we run the model two times which generates two groups of different prediction results correspondingly. After calculation, the % error is presented on Figure 4-8c by taking the difference between the predicted value and the measured ground truth dived by the ground truth times 100%. We found that all the predicted data has a percentage error smaller than 25%. Specifically, we found most of the errors lie in between 10% with a slightly higher percentage error for fill factor. This aligns well with our physical understanding as Fill Factor largely depends on the solar cell fabrication process and intralayer film quality which might not be perfectly reflected by the %T, SRPL and TRPL spectrum.

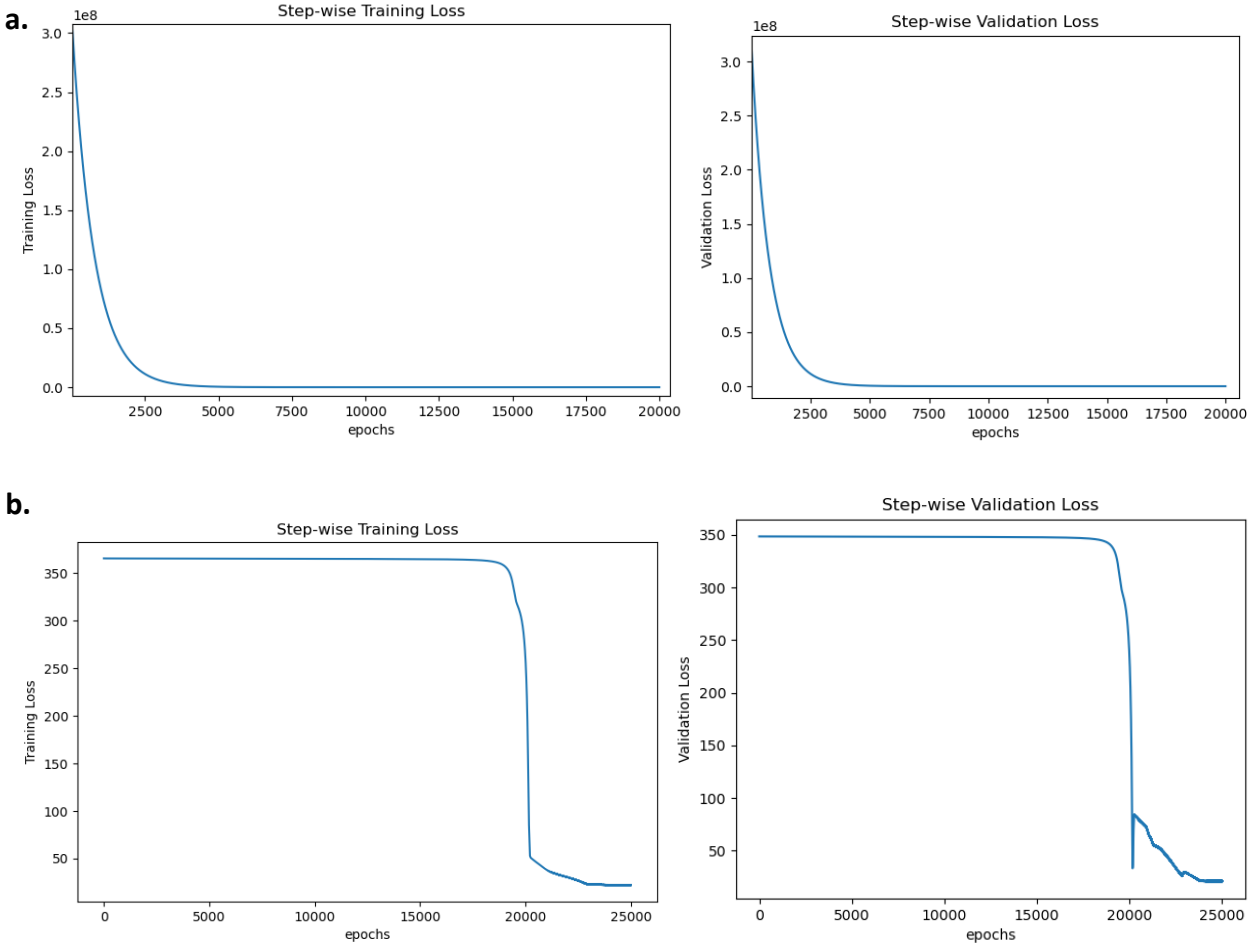


Figure 4-7 Stepwise training loss during training (left) and validation (right) forward process for 20 input samples a. training loss based on l2-norm feedback cost function, b. training loss based on l1-norm feedback cost function.

a.

Measured Output (Ground Truth)		
Voc	Jsc	FF
969.8	24.98	77.9
965.8	24.7	79
899.9	24.59	66.1
946.1	24.85	78.1

b.

```
valid is tensor([[1047.7365,   26.0481,   69.3757],
        [1024.7974,   25.4977,   67.8711],
        [ 884.1210,   22.1275,   58.6222],
        [1004.8279,   25.0182,   66.5596]], grad_fn=<AddmmBackward>)
```

c.

```
valid is tensor([[1054.3033,   24.2443,   67.5757],
        [1060.6600,   24.3868,   67.9813],
        [1064.8439,   24.4799,   68.2157],
        [1062.0492,   24.4180,   68.0700]], grad_fn=<AddmmBackward>)
```

d.

% Error		
Voc	Jsc	FF
8.03	4.72	10.94
6.11	3.23	11.56
1.75	10.01	11.31
6.21	0.68	14.78

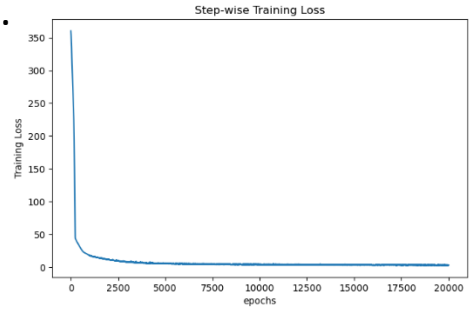
e.

% Error		
Voc	Jsc	FF
8.71	2.15	13.25
9.82	1.27	13.95
18.33	0.45	3.2
12.26	1.74	12.84

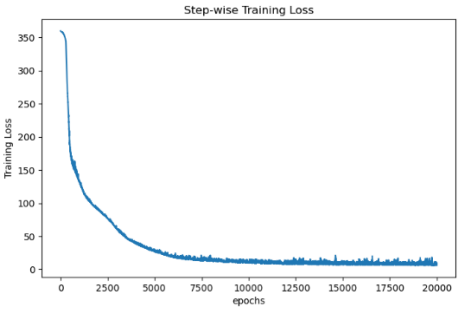
Figure 4-8 a. Measured validation dataset output as ground truth. b, c. two model predicted values running two different training seeds d, e. percent error based on trial b and c, correspondingly.

Similar prediction is carried out for the larger dataset. After analyzing the data distribution histogram of the Voc, Jsc and FF output values, we found a small fraction of data act as the outlier which may influence the prediction results. With most of the data distributed in the center range of the gaussian distribution, those outliers tend to have an extremely different output value (either Voc, Jsc or FF) comparing to the rest of the data. If the validation dataset includes any of those outliers, the predicted result of that specific sample tends to have extremely large variance from the ground truth. Apparently, as 80% of the input data lies in a reasonable range, due to the lack of data for the outliers, the model has insufficient data to learn those outliers' trends. Hence, we remove the outliers from the validation dataset and carry out the prediction. A better prediction error is found for the larger dataset owe to the large capacitance of the input data as shown in Figure 4-9.

a.



b.



c.

Test Sets					Prediction Results			Percent Error		
#	Voc (mV)	Jsc (mA/cm2)	FF	PCE	Voc (mV)	Jsc (mA/cm2)	FF	Voc (mV)	Jsc (mA/cm2)	FF
1.5.3	1029.5	24.87	58.2	14.9	993.81	24.40	72.34	3.47	1.89	24.30
1.5.4	989.1	24.69	48.6	11.87	992.32	24.37	72.24	0.33	1.30	48.64
2.9.1	1078.3	23.96	73.3	18.94	1016.69	24.84	73.85	5.71	3.67	0.75
3.7.1	1099.1	23.9	77.2	20.28	1088.99	26.22	78.63	0.92	9.71	1.85
3.7.2	1090.2	23.7	78	20.16	1064.00	25.74	76.98	2.40	8.61	1.31
5.7.2	889.9	23.72	81.2	17.14	1018.81	24.88	73.99	14.49	4.89	8.88
5.7.3	901.1	24.29	81.7	17.88	990.75	24.34	72.14	9.95	0.21	11.70
6.5.1	911.3	24.11	69.5	15.27	1018.98	24.88	74.01	11.82	3.19	6.49
								Avg		
								6.13	4.18	12.99

Figure 4-9 a. Stepwise training loss for training dataset and b. validation dataset based on 212 input samples c. Ground truth output dataset with predicted results from model and percent error.

As the results have presented, we randomly choose 212 and 8 datapoints from the optimized dataset as training data and validation data, correspondingly. Among all the predicted output, around 30% of the predicted results have a %error of 10%, and 65% of the predicted value have a %error less than 5%. The further decrease of percent error comparing to the first smaller dataset takes credit to the enlarge of learning input dataset which facilitates the training of the model.

4.5 Linkage between Physical Science and Machine-learning Algorithm

However, even though neural network takes a good prediction and is capable of learning the possible relationships inside the Blackbox between the input and output, it is still interested to learn how much weights do the three input curves count towards the final predicted results. Linear/ Polynomial Regression algorithm can fit those different inputs into a linear or polynomial relation with the output through predicting the parameters corresponding to each input, acting as the weight of each input. Instead of merely loading the whole three graphs into the algorithm, we also manually select some characterization points as input from the experimentally acquired raw data. For the percent transmission curve, we find 8 specific values from the spectrum including: the local minimum wavelength & local minimum percent transmission intensity; the percent transmission at wavelength of 550nm and 850nm; the average transmission between wavelength range in 750-800nm; the log fitting equation of the transmission curve $y = Ae^{Bx} + C$.

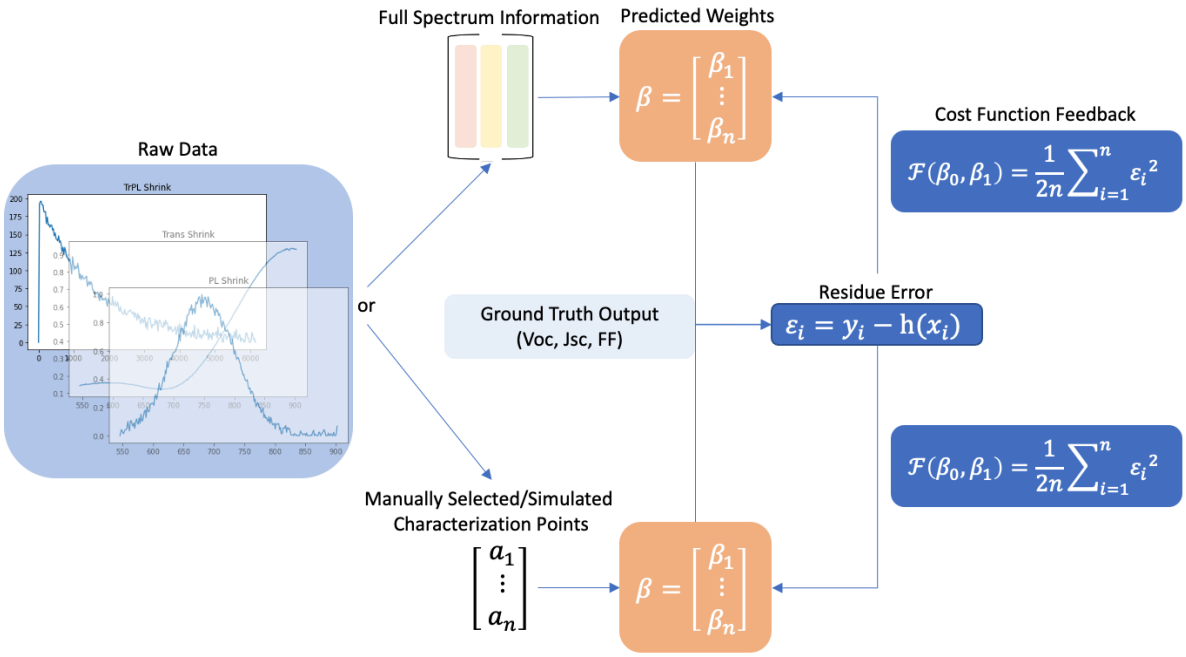


Figure 4-10 Architecture of Linear/ Polynomial Regression Algorithm.

Similarly, 11 parameters are extracted from the SRPL spectrum: PL peak starting wavelength and intensity; PL peak ending wavelength and intensity; maximum peak position wavelength and intensity; Full-width-half-maximum (FWHM) of the PL peak; two linearly fitted equations y=Ax± B for the peak increasing slope and the decreasing slope. For the TRPL spectrum, we extract the maximum TRPL intensity corresponding to the time; the lifetime of the carriers (one-third of the maximum intensity of the curve); a sixth-order polynomial fitting curve equation $y=Ax^6+Bx^5+Cx^4+Dx^3+Ex^2+Fx+G$. Note that all the fittings are carried out through python converging packages scipy.optimize and numpy.linalg. After preparing all the data, the architecture of the regression algorithm is presented in Figure 4-10. Either the raw 750 data or the in total 29 characteristic points could be put into the algorithm. A basic linear/polynomial follow the math of $y_i = \beta_0 + \beta_i x_i + \varepsilon_i = \mathrm{h}(x_i) + \varepsilon_i$ where x is the input matrix, y is the output matrix, h(x) is the predicted relationship between x and y as β representing the weights of each input and

ε as residue error. The square error cost function is defined as $\mathcal{F}(\beta_0, \beta_1) = \frac{1}{2n}\sum_{i=1}^{n} {\varepsilon_i}^2$ as the algorithm find the β_0, β_1 such that $\mathcal{F}(\beta_0, \beta_1)$ is minimized. The predicted $\beta_1 = \frac{SS_{xy}}{SSxx}$, $\beta_0 = \bar{y} - \beta_1\bar{x}$ follow the above relationship where $SS_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$ and$SS_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} {x_i}^2 - n\bar{x}^2$. As the algorithm has optimized the parameters of the equation after training, we are able to predict any results with the validation dataset. Figure 4-11 shows the prediction results based on the linear regression algorithm with all 750 spectrum data points as training inputs. Need to note that linear/ polynomial regression predicts one output at a time instead of an output matrix which requires us to run the model three times based on the three demanded photovoltaic performance values. As Figure 4-11 b – d has presented, a relatively similar percent error is produced compared with the neural network algorithm which proves the robust dataset and algorithm we have developed.

a.

```
Valdiation Output:
[[1029.5    24.87   58.2 ]
 [ 989.1    24.69   48.6 ]
 [1078.3    23.96   73.3 ]
 [1099.1    23.9    77.2 ]
 [1090.2    23.7    78.  ]
 [ 889.9    23.72   81.2 ]
 [ 901.1    24.29   81.7 ]
 [ 911.3    24.11   69.5 ]]
```

b.

```
predicted response:
[1015.92575815 1110.07903385 1143.75076494 1011.87789808  955.66407599
  960.39535699  794.17107368  976.98423756]
[ -1.31852762  12.23122372   6.06981034  -7.9357749  -12.34048101
   7.92171671 -11.86648833   7.2077513 ]
Percent error is: 8.361471741626925 %
```

c.

```
predicted response:
[26.34973984 25.28446716 24.29249165 24.01294972 23.97860954 26.44943292
 25.23407802 24.8027731 ]
[ 5.94989882  2.40772444  1.38769469  0.47259299  1.17556769 11.50688414
  3.88669421  2.87338491]
Percent error is: 3.7075552364101614 %
```

d.

```
predicted response:
[59.98097412 67.85968802 48.48805404 69.77019148 63.62891276 68.88027213
 64.6017628  75.73907421]
[  3.06009299  39.6289877  -33.84985807  -9.62410431 -18.42447082
 -15.17207866 -20.9280749    8.97708519]
Percent error is: 18.708094078326013 %
```

Figure 4-11 a. Ground truth validation output. Predicted response and %error of b. Voc, c. Jsc and d. FF.

	Parameter			Parameter	
Transmission	a_1	Local min wavelength ג		a_{16}	$y = a_{16}x + a_{17}$
	a_2	Local min %Trans		a_{17}	
	a_3	%T @ 550 nm		a_{18}	$y = a_{18}x + a_{19}$
	a_4	%T @ 850 nm		a_{19}	
	a_5	Avg %T bw750-800 nm	Time-Resolved PL	a_{20}	Max TrPL Lifetime
	a_6	$y = a_6 e^{a_7 x} + a_8$		a_{21}	Max TrPL Intens.
	a_7			a_{22}	1/3 of max Lifetime
	a_8			a_{23}	

Spectrum-Resolved PL	a_9	Peak Start ג		a_{24}	$y = a_{23}x^6 + a_{24}x^5 + a_{25}x^4 + a_{26}x^3 + a_{27}x^2 + a_{28}x + a_{29}$
	a_{10}	Peak Maximum ג		a_{25}	
	a_{11}	Peak End ג		a_{26}	
	a_{12}	Peak Start Intens.		a_{27}	
	a_{13}	Peak Maximum Intens.		a_{28}	
	a_{14}	Peak End Intens.		a_{29}	
	a_{15}	FWHM			

Table 2 29 characteristic input points specifications.

Additionally, 29 characteristic values are also used to train the model and predict the weighting parameter correspondingly. Table 2 provides a specification of each parameter as introduced at the beginning of this section. Through linear regression algorithm, we found that the fitted curve trend parameters (e.g., $a_{6\text{-}8}$, a_{16}, a_{18}, $a_{26\text{-}28}$) have a stronger influence on the training algorithm with other parameters lightly weighted. On the other hand, some parameters have almost no weights contributing to the algorithm like a_1, and a_2 (Figure 4-12). This aligns with our fundamental physical understanding as the shape of those three graphs reflects the bandgap, absorption, and film property information of the cell while the local minimum always happens outside the valuable wavelength range. Surprisingly, under this training result, we find little influence from the TRPL lifetime parameter on the algorithm (a_{22}). We suppose this is due

to the independence between each input parameter as this might not perfectly reflect the truth for a solar cell device. To solve this problem, the quadratic form of these 29 characteristic points is developed (Figure 4-13).

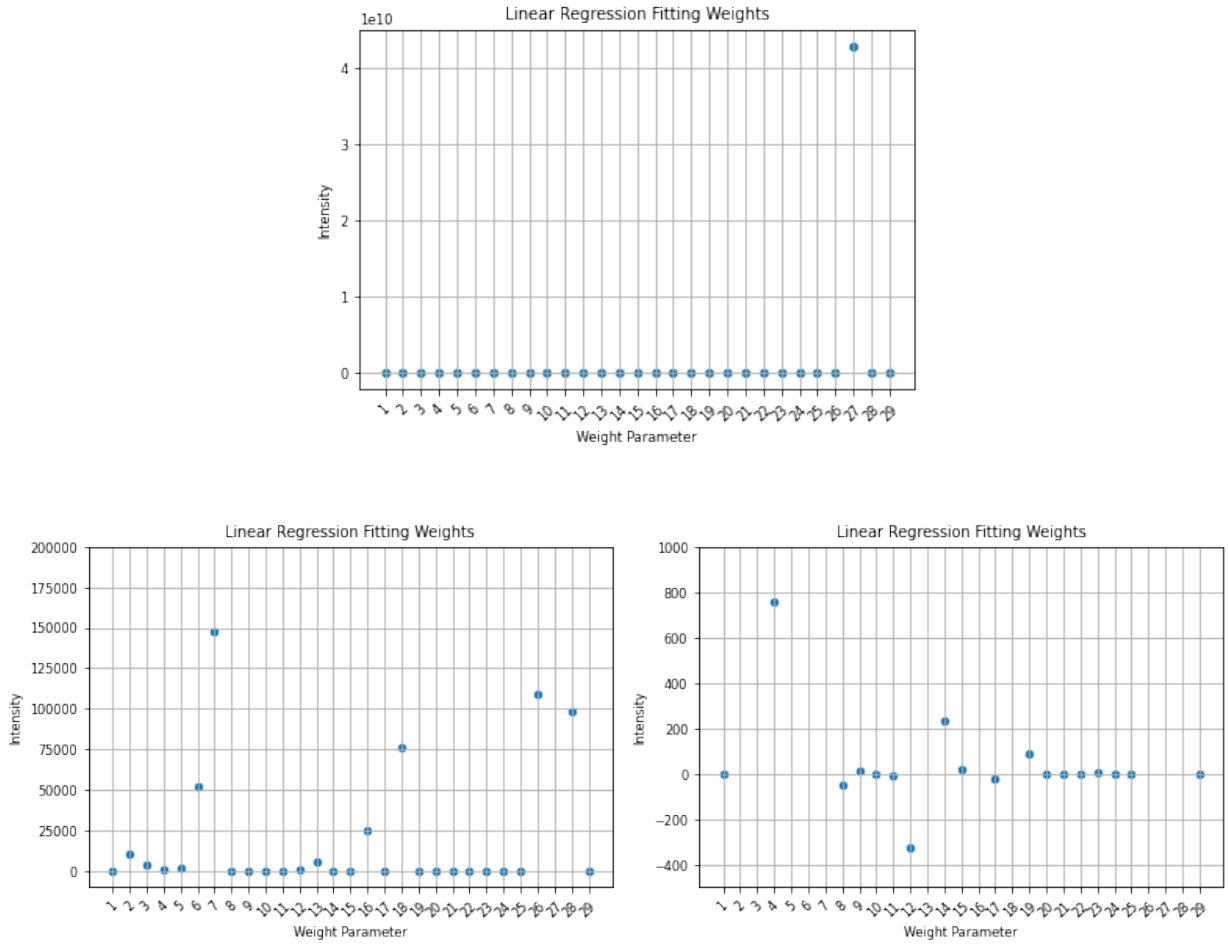


Figure 4-12 a. Weighted parameter intensity corresponds to the linear regression prediction model. b. c. Further zoomed in version of weighting intensity from a.

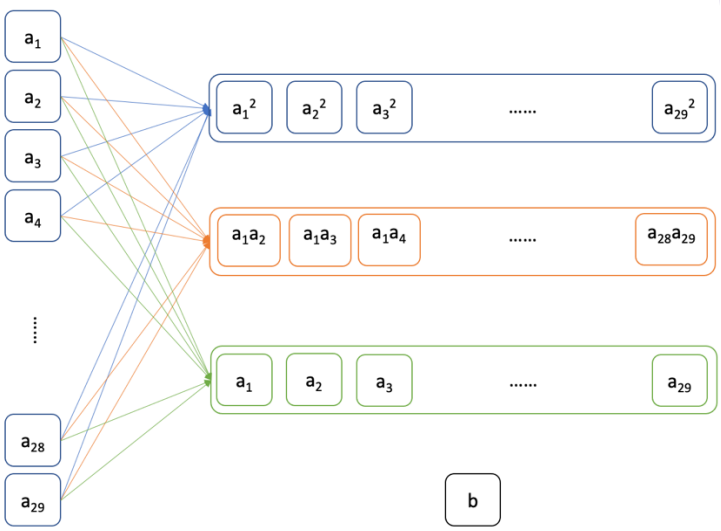


Figure 4-13 Quadratic form of the 29 characteristic inputs.

The 29 points are extended into 0^{th} order term, 1^{st} order term, and 2^{nd} order terms to intercorrelate with each other. With the quadratic extension input, we have again put all the quadratic forms into a polynomial regression algorithm. Here, the predicted weights corresponding to each quadratic parameter is shown in Figure 4-14. Note only the right top part of the graph matters with all the rest value equal to 0. Each small square represents the predicted parameter that correlates to the input value on that column and row. E.g., the (1, a_1) square represents the predicted parameter (weights) corresponding to the input value of a_1. The diagonal columns in the heat map represents the weights in front of the a_1^2, a_2^2, … a_{29}^2 term, correspondingly. For aesthetic reasons, the single heat map block located at the bottom left corner of the graph represents the weights corresponding to the 0^{th} order term, b. This leads to more reasonable results showing an increasing weight corresponding to the TRPL lifetime (a_{20}, a_{21}, a_{22}) while remaining the influence of the similar weights based on the other fitting curve parameters. With these sections, we are able to learn the potential matches between the device physics and the prediction algorithms.

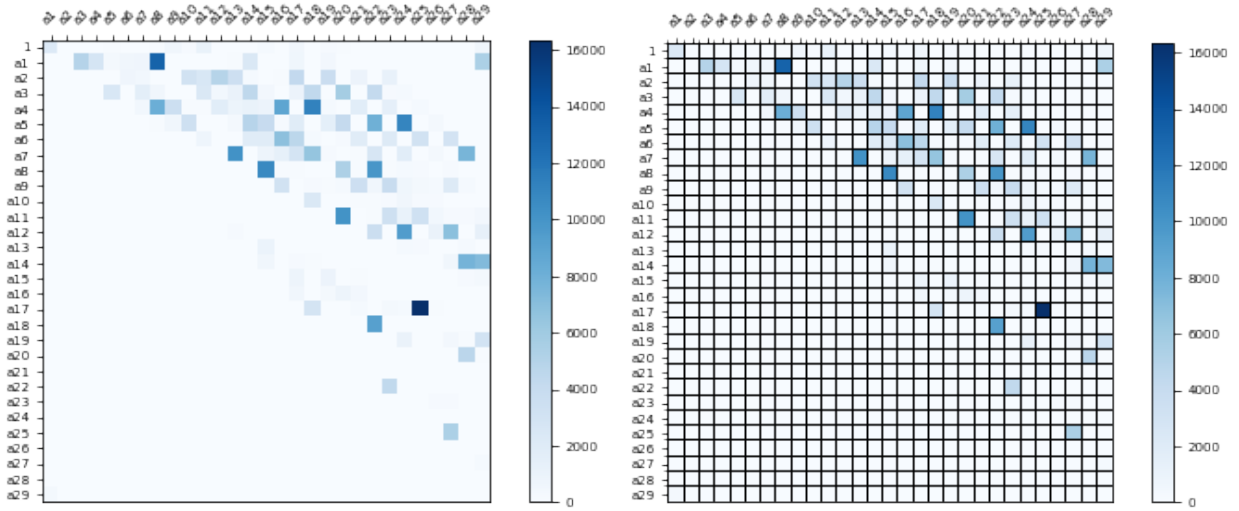


Figure 4-14 Quadratic weights of the 29 characteristic inputs. Note the bottom left corner represents the 0th order b term. The rest of bottom left triangle is supposed to leave blank. Right side is an exact same graph as left only with an addition of gridline to b.

Figure 4-15. is presented as a supportive material showing that the algorithm learns well based on the training data vs. predicted data difference through the input. Figure 4-15a proves that for each input quadratic datapoint, the training predictions have a small %error for all the datapoints smaller than 5%. Figure 4-15b presents the scatter plot showing the distribution of the ground truth data and the predicted values.

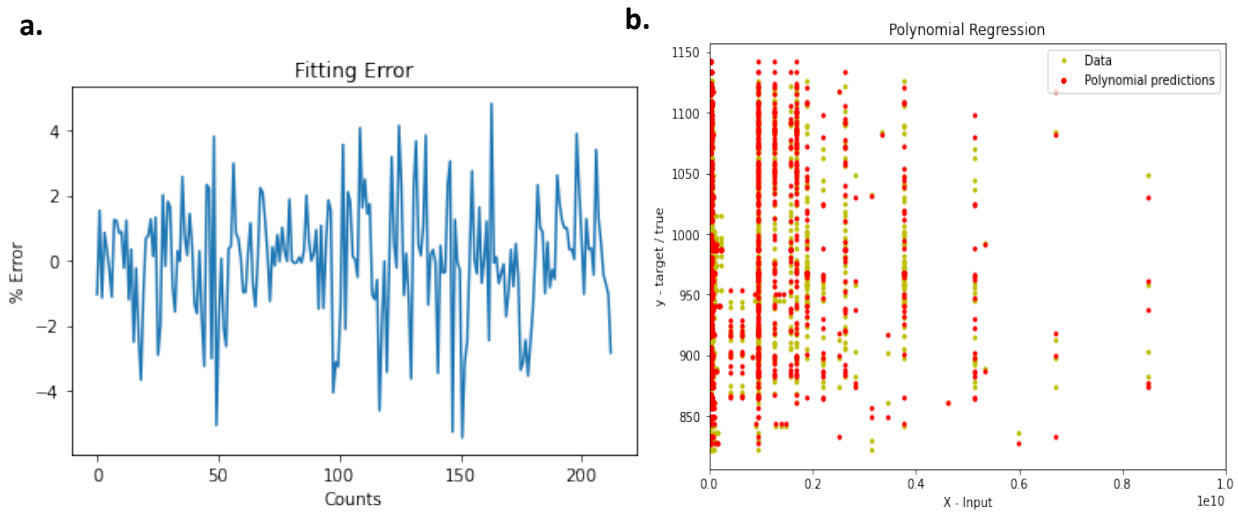


Figure 4-15 a. %Error between each quadratic input data and model-predicted data b. Scatter plot of polynomial regression predicted points with the raw input data.

In all, we have successfully developed three algorithms predicting the photovoltaic performance of the solar cell from the percent transmission, spectrum-resolved photoluminescence, and time-resolve photoluminescence spectrums. For a randomly selected validation dataset, we could reach the percent prediction error lies under 10% with more than 60% of the validation output within 5% error. The best prediction error we have reached is 0.21% error. The consistency of the prediction results validates the algorithm and dataset is accurate and robust. There are still plenty of interesting staff we are considering carrying on including developing more machine-learning algorithm (Decision tree, Random Forest, Support Vector Machines (SVM), etc.); investigating the influence of sample degradation to the prediction model; further optimize the dataset and algorithm pursuing an even smaller %error as well as generate a possible classifier algorithm which can predict multiple different structures.

Chapter 5 Conclusions and Future Directions

5.1 Conclusion

There are three main conclusions in approaching novel perovskites photovoltaic devices through Machine Learning and Interfacial Engineering.

First, TcTa acts as a promising candidate in preventing sputter damage while maintaining high intralayer carrier mobilities for a p-i-n diode and metal contact interface. Under the scalable baseline ITO/SnO_2/(MAFA)Pb$(IBrCl)_3$/PV2000/Au structure, with the addition of the TcTa layer, the perovskite solar cell tends to have a drop of PCE with less than 1%. 3 nm is the threshold thickness of the buffer layer in maintaining the sample surface integrity from sputter plasma damage while enabling carrier tunneling effects. The thicker the buffer layer, the lower Jsc and FF show for a standard light JV-sweep. Nickle works as a good substitution in getting rid of the commercially inappropriate Au-electrode due to its mechanical toughness, high electricity conductivity, and low contact resistance. With the help of a 3 nm TcTa buffer layer, a TcTa/Ni-based solar cell is fabricated with a PCE of 10.69%.

Second, to predict the photovoltaic properties of a 3D/2D perovskite solar cell, SRPL, TRPL, and % Transmission spectrums are collected to feed a well-designed neural network machine learning algorithm. After training the algorithm using more than 200 data points, with a stepwise learning rate of 0.0001 and 20000 learning epochs, a less than 5% regression prediction

error is reached for more than 65% of all the predicting results. The best prediction error is reached for less than 1% with a total of 90% of the predicting result having an error less than 10%.

Last, linear regression and polynomial regression algorithms are developed to investigate the possible relationship between physical science parameters and the machine learning algorithm. After analyzing the weights of each training parameter, the parameters representing the shape information of the three spectrums tend to have a comparably large weight that the machine takes more into consideration during the training process. Except for the spectrum shape parameters, most of the secondly weighted parameters lie with our physical understanding including the lifetime of the cell, the SRPL intensity at the wavelength corresponding to the perovskite bandgap wavelength. Further investigations will be carried out in the future.

5.2 Future Works

Based on the current work, more research will be carried out in machine learning parts. Data including degraded solar cells under the same structure will be considered in the algorithm. The neural network algorithm will be further modified to fit the new dataset. Besides, new parameters corresponding to the three input spectrums will be obtained without the influence of curve shape information. A further investigation on the weights of the pure physical-based inputs will be studied to better understand the possible physical meaning of the algorithm.

To put in a big picture, machine-learning induced regression algorithm in predicting semiconductor device performance not limited to perovskite could be further investigated. Possible

future directions include predicting the photovoltaic responses of solar cells based on various solar cell structures (with the addition or subtraction of layers, etc.); predicting the device's photovoltaic performances when discovering various new materials; predicting LEDs or photodetector's performances based on various single film properties, etc. In all, in the era of highly developed information technology, machine learning will be an irreplaceable technique in next-generation semiconductor studies.

A. List of Materials, Equipment, Vendors, and Methods

A.1 Materials and Methods

Perovskite Solar Cell fabrication procedure follows the bottom-up sequence of the solar cell: pure glass, bottom electrode (ITO), Electron transport layer (SnO_2/TiO_2), perovskite ($(MAFACs)Pb(BrICl)_3$), hole transport layer (PV2001), a buffer layer (LiF, MoO_3, TcTa), top electrode (Ni, Al, Au, Ag). The baseline thin film fabrication process follows Richard Swartout's Ph.D. thesis. LiF and MoO_3 are thermally evaporated with the source material purchased from Lumtech Inc. TcTa, ZnPc, and other large organic molecules are also thermally evaporated source from Lumtech. Inc. Ni, Al, Au is sputtered for 70nm under an Argon environment.

A.2 Equipment, Facilities and Locations

The equipment is mainly located and utilized from MIT ONE Lab and MIT.Nano with additional help from MRL in MIT Building 13. The detailed equipment is listed below:

Core Fabrication facilities: Glove Box (ONE Lab), Dry Box (ONE Lab), Spin Coater (ONE Lab), Thermal Evaporator (ONE Lab), Sputter System (ONE Lab), Slot-die coater (MIT.Nano), etc.

Core Characterization facilities: Gemini Scanning Electron Microscope (SEM, MIT.Nano), Smart lab X-ray Diffraction (XRD, MRL), small angle XRD (MRL), Time-resolved Photoluminescence (TrPL, ONE LAB), Dektak XD/150 (MIT.Nano), Lifetime measurement wave lab (ONE Lab), JV-Curve Power Meter (ONE Lab), ellipsometer (ONE Lab), etc.

B. Supplementary Information of Chapter 3 and Chapter 4 (Appendix)

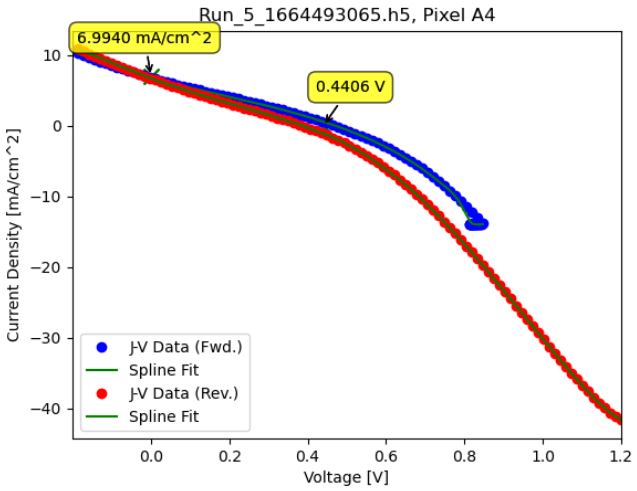


Figure B-1 JV-sweep of Sputtered Ni Solar Cell.

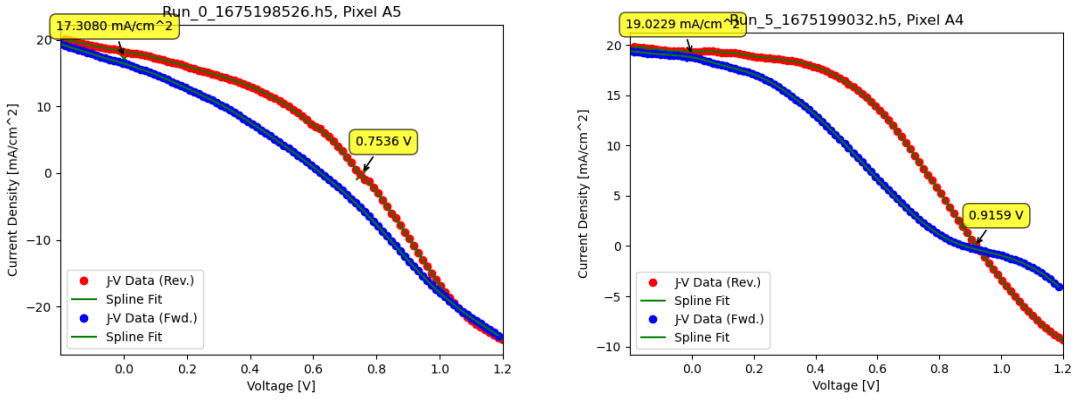


Figure B-2 E-beam evaporated Ni Solar Cell.

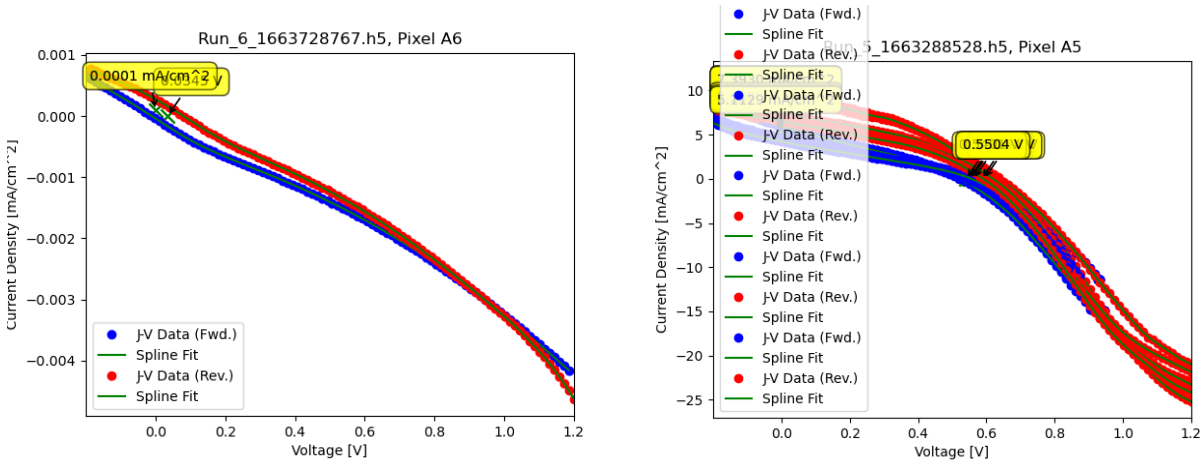


Figure B-3 Thermally evaporated 10nm LiF + thermally evaporated 70nm Au Solar Cell.

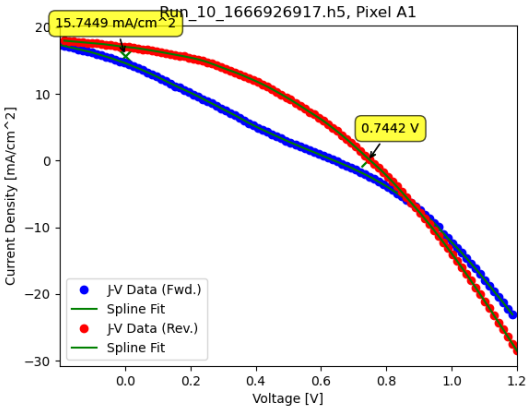


Figure B-4 Thermally evaporated 10nm MoO3 + sputtered 70nm Al Solar Cell.

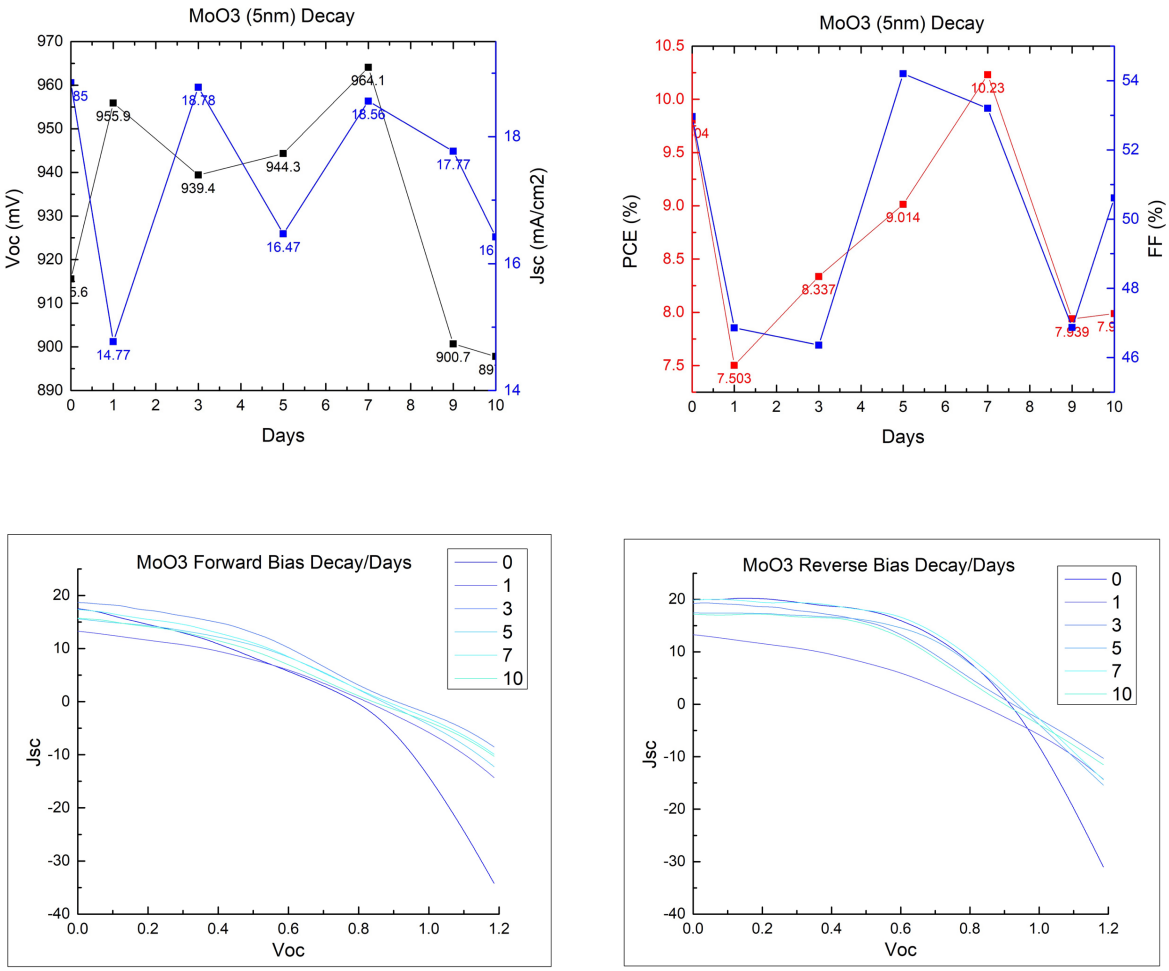


Figure B-5. MoO3 + Ni Decay Investigation.

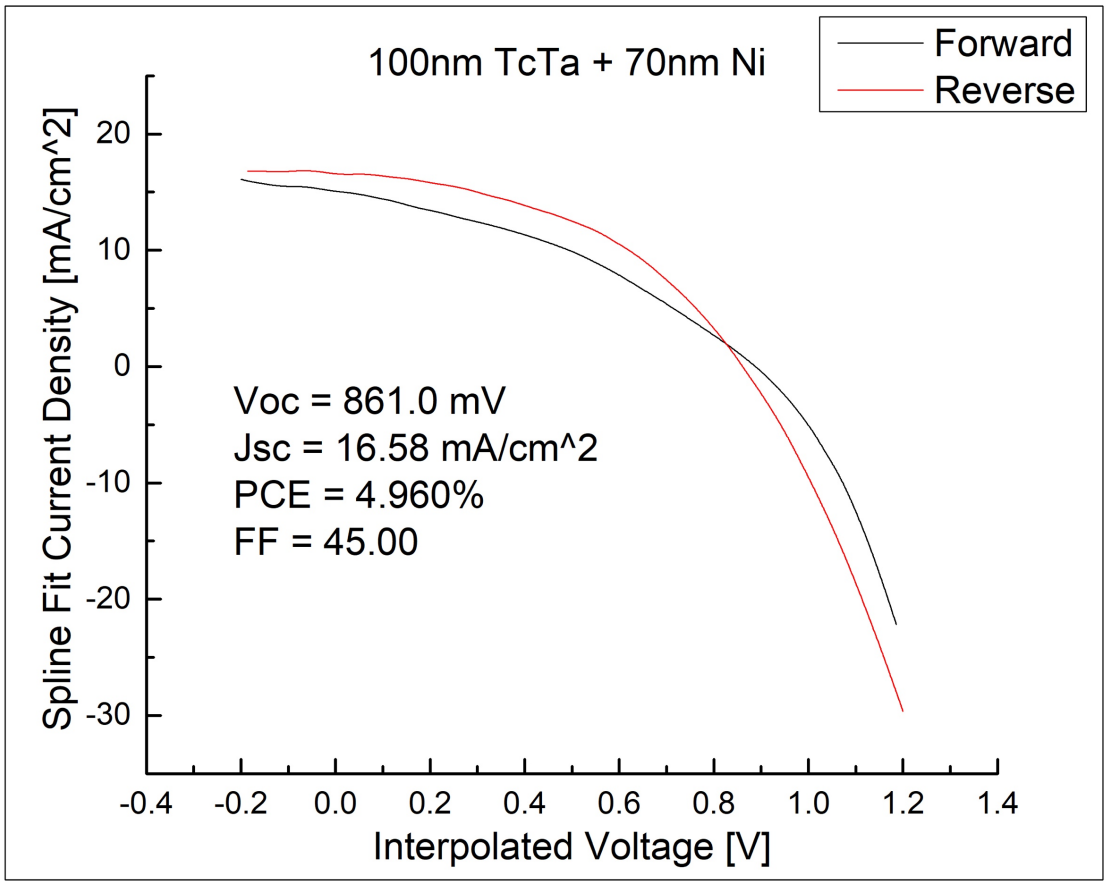


Figure B-6 TcTa as HTL material.

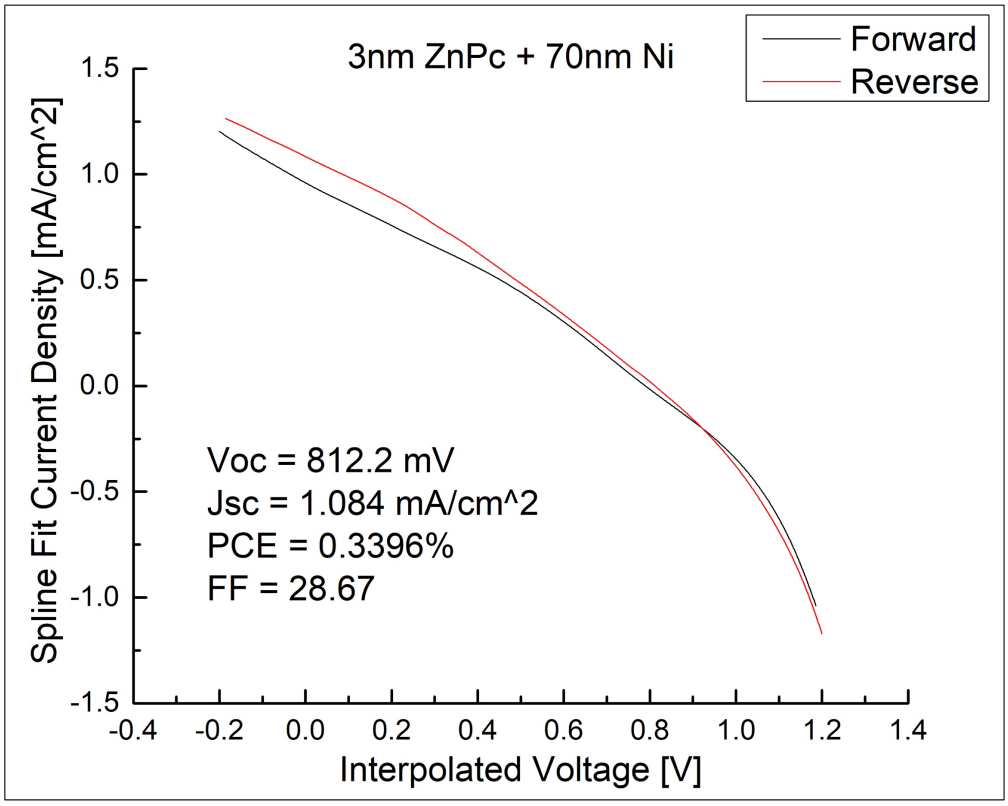


Figure B-7 3nm ZnPc buffer layer reference cell.

References

1. Leng, K., et al., *From bulk to molecularly thin hybrid perovskites.* Nature Reviews Materials, 2020. **5**(7): p. 482-500.
2. Kim, H.-S., et al., *Ferroelectric Polarization in CH3NH3PbI3 Perovskite.* The Journal of Physical Chemistry Letters, 2015. **6**(9): p. 1729-1735.
3. Stranks, S.D. and H.J. Snaith, *Metal-halide perovskites for photovoltaic and light-emitting devices.* Nature Nanotechnology, 2015. **10**(5): p. 391-402.
4. Si, H., et al., *Emerging Conductive Atomic Force Microscopy for Metal Halide Perovskite Materials and Solar Cells.* Advanced Energy Materials, 2020. **10**(10): p. 1903922.
5. Lei, Y., et al., *Perovskite superlattices with efficient carrier dynamics.* Nature, 2022. **608**(7922): p. 317-323.
6. Wang, F., et al., *Defects engineering for high-performance perovskite solar cells.* npj Flexible Electronics, 2018. **2**(1): p. 22.
7. Chen, Y., et al., *Strain engineering and epitaxial stabilization of halide perovskites.* Nature, 2020. **577**(7789): p. 209-215.
8. Lei, Y., et al., *A fabrication process for flexible single-crystal perovskite devices.* Nature, 2020. **583**(7818): p. 790-795.
9. Fan, Z., et al., *Ferroelectricity of CH3NH3PbI3 Perovskite.* The Journal of Physical Chemistry Letters, 2015. **6**(7): p. 1155-1161.
10. Lu, Y., et al., *Rational Design of A Chemical Bath Deposition Based Tin Oxide Electron Transport Layer for Perovskite Photovoltaics.* Advanced Materials. **n/a**(n/a): p. 2304168.
11. Gagnon, P., Margolis, Robert, Melius, Jennifer, Phillips, Caleb, and Elmore, Ryan, *Rooftop Solar Photovoltaic Technical Potential in the United States: A Detailed Assessment.* United States N. p., 2016.
12. Moon, S., et al., *Highly efficient single-junction GaAs thin-film solar cell on flexible substrate.* Scientific Reports, 2016. **6**(1): p. 30107.
13. Saga, T., *Advances in crystalline silicon solar cell technology for industrial mass production.* NPG Asia Materials, 2010. **2**(3): p. 96-102.
14. Lei, Y., Y. Chen, and S. Xu, *Single-crystal halide perovskites: Opportunities and challenges.* Matter, 2021. **4**(7): p. 2266-2308.
15. Zhang, W., G.E. Eperon, and H.J. Snaith, *Metal halide perovskites for energy applications.* Nature Energy, 2016. **1**(6): p. 16048.
16. Huang, C., et al., *Facile fabrication of highly efficient ETL-free perovskite solar cells with 20% efficiency by defect passivation and interface engineering.* Chemical Communications, 2019. **55**(19): p. 2777-2780.
17. Donlan, E.A., et al., *Compositional nanodomain formation in hybrid formate perovskites.* Chemical Communications, 2017. **53**(81): p. 11233-11236.
18. Lu, C., et al., *Demystifying phase transformations in metal halide perovskites.* Matter, 2021. **4**(8): p. 2627-2629.
19. Foley, B.J., et al., *Temperature dependent energy levels of methylammonium lead iodide perovskite.* Applied Physics Letters, 2015. **106**(24).

20. Wehrenfennig, C., et al., *Charge carrier recombination channels in the low-temperature phase of organic-inorganic lead halide perovskite thin films.* APL Materials, 2014. **2**(8).
21. Saidaminov, M.I., et al., *Planar-integrated single-crystalline perovskite photodetectors.* Nature Communications, 2015. **6**(1): p. 8724.
22. Xu, W., et al., *Rational molecular passivation for high-performance perovskite light-emitting diodes.* Nature Photonics, 2019. **13**(6): p. 418-424.
23. Park, J., et al., *Controlled growth of perovskite layers with volatile alkylammonium chlorides.* Nature, 2023. **616**(7958): p. 724-730.
24. Lekina, Y., et al., *The effect of organic cation dynamics on the optical properties in (PEA)2(MA)[Pb2I7] perovskite dimorphs.* Journal of Materials Chemistry C, 2021. **9**(47): p. 17050-17060.
25. Fu, Y., et al., *Cation Engineering in Two-Dimensional Ruddlesden–Popper Lead Iodide Perovskites with Mixed Large A-Site Cations in the Cages.* Journal of the American Chemical Society, 2020. **142**(8): p. 4008-4021.
26. Osada, M. and T. Sasaki, *Nanoarchitectonics in dielectric/ferroelectric layered perovskites: from bulk 3D systems to 2D nanosheets.* Dalton Transactions, 2018. **47**(9): p. 2841-2851.
27. Shao, Y., et al., *Grain boundary dominated ion migration in polycrystalline organic–inorganic halide perovskite films.* Energy & Environmental Science, 2016. **9**(5): p. 1752-1759.
28. Ashurov, N., et al., *Current state and perspectives for organo-halide perovskite solar cells. Part 1. Crystal structures and thin film formation, morphology, processing, degradation, stability improvement by carbon nanotubes. A review.* Modern Electronic Materials, 2017. **3**(1): p. 1-25.
29. Yoo, J.J., et al., *Efficient perovskite solar cells via improved carrier management.* Nature, 2021. **590**(7847): p. 587-593.
30. Wang, L., et al., *Non-precious transition metals as counter electrode of perovskite solar cells.* Energy Storage Materials, 2017. **7**: p. 40-47.
31. Svanström, S., et al., *Degradation Mechanism of Silver Metal Deposited on Lead Halide Perovskites.* ACS Applied Materials & Interfaces, 2020. **12**(6): p. 7212-7221.
32. Li, J., et al., *Direct Evidence of Ion Diffusion for the Silver-Electrode-Induced Thermal Degradation of Inverted Perovskite Solar Cells.* Advanced Energy Materials, 2017. **7**(14): p. 1602922.
33. Zhao, J., et al., *Is Cu a stable electrode material in hybrid perovskite solar cells for a 30-year lifetime?* Energy & Environmental Science, 2016. **9**(12): p. 3650-3656.
34. Kaltenbrunner, M., et al., *Flexible high power-per-weight perovskite solar cells with chromium oxide–metal contacts for improved stability in air.* Nature Materials, 2015. **14**(10): p. 1032-1039.
35. Abdy, H., et al., *Investigation of metal-nickel oxide contacts used for perovskite solar cell.* AIP Advances, 2019. **9**(1): p. 015216.
36. Abdollahi Nejand, B., V. Ahmadi, and H.R. Shahverdi, *New Physical Deposition Approach for Low Cost Inorganic Hole Transport Layer in Normal Architecture of Durable Perovskite Solar Cells.* ACS Applied Materials & Interfaces, 2015. **7**(39): p. 21807-21818.
37. Jiang, Q., et al., *Nickel-Cathoded Perovskite Solar Cells.* The Journal of Physical Chemistry C, 2014. **118**(45): p. 25878-25883.

38. Jeong, I., et al., *Highly efficient perovskite solar cells based on mechanically durable molybdenum cathode.* Nano Energy, 2015. **17**: p. 131-139.
39. Tavakoli, M.M., et al., *A Dopant-Free Hole Transporting Layer for Efficient and Stable Planar Perovskite Solar Cells.* physica status solidi (RRL) – Rapid Research Letters, 2020. **14**(7): p. 2000147.
40. Ofuonye, B., et al., *Electrical and microstructural properties of thermally annealed Ni/Au and Ni/Pt/Au Schottky contacts on AlGaN/GaN heterostructures.* Semiconductor Science and Technology, 2014. **29**(9): p. 095005.
41. Liu, Y., et al., *How Machine Learning Predicts and Explains the Performance of Perovskite Solar Cells.* Solar RRL, 2022. **6**(6): p. 2101100.
42. Liu, Y., et al., *Machine Learning for Perovskite Solar Cells and Component Materials: Key Technologies and Prospects.* Advanced Functional Materials, 2023. **33**(17): p. 2214271.
43. Srivastava, M., et al., *Machine Learning Enables Prediction of Halide Perovskites' Optical Behavior with >90% Accuracy.* ACS Energy Letters, 2023. **8**(4): p. 1716-1722.
44. Tao, Q., et al., *Machine learning for perovskite materials design and discovery.* npj Computational Materials, 2021. **7**(1): p. 23.
45. Zhao, Y., et al., *Discovery of temperature-induced stability reversal in perovskites using high-throughput robotic learning.* Nature Communications, 2021. **12**(1): p. 2191.
46. Hartono, N.T.P., et al., *How machine learning can help select capping layers to suppress perovskite degradation.* Nature Communications, 2020. **11**(1): p. 4172.
47. Ma, T., et al., *Unveil the Full Potential of Integrated-Back-Contact Perovskite Solar Cells Using Numerical Simulation.* ACS Applied Energy Materials, 2018. **1**(3): p. 970-975.
48. Srivastava, M., et al., *Machine Learning Roadmap for Perovskite Photovoltaics.* The Journal of Physical Chemistry Letters, 2021. **12**(32): p. 7866-7877.
49. Brenes, R., et al., *Benefit from Photon Recycling at the Maximum-Power Point of State-of-the-Art Perovskite Solar Cells.* Physical Review Applied, 2019. **12**(1): p. 014017.