Autonomous Materials Discovery for Organic Photovoltaics

Changhyun Hwang
University of Illinois, Urbana-Champaign
chwang12@illinois.edu

Seungjoo Yi
University of Illinois, Urbana-Champaign
yi22@illinois.edu

David M. Friday
University of Illinois, Urbana-Champaign
friday1@illinois.edu

Nicholas H. Angello
University of Illinois, Urbana-Champaign
angello3@illinois.edu

Tiara C. Torres-Flores
University of Illinois, Urbana-Champaign
tiaract2@illinois.edu

Nick Jackson
University of Illinois, Urbana-Champaign
jacksonn@illinois.edu

Martin D. Burke
University of Illinois, Urbana-Champaign
mdburke@illinois.edu

Charles M. Schroeder
University of Illinois, Urbana-Champaign
cms@illinois.edu

Ying Diao
University of Illinois, Urbana-Champaign
yingdiao@illinois.edu

Abstract

We aim to develop an AI-guided autonomous materials design approach to discover high-performance organic photovoltaics (OPVs). Autonomous synthesis, automated characterization, and AI-based methods will be integrated into a closed-loop approach to drive molecular discovery guided by target criteria for OPV performance: efficiency and stability. The long-term goal of the project is two-fold: (1)in terms of fundamental science, we aim to fill key knowledge gaps in understanding how molecular structure determines OPV stability and efficiency, and advance the science of closed-loop autonomous discovery by learning how to synergistically integrate AI, automated synthesis, and automated testing. (2)In terms of technology, we aim to meet the "10-10" target (10% efficiency and 10-year stability for OPV materials) to make OPVs a commercial reality for next-generation energy capture applications and for mitigating climate change.

1 Overview

Organic photovoltaics (OPVs) are next-generation devices for harvesting renewable energy that are becomingly increasingly desired over traditional silicon solar cells due to lighter weight, lower cost, increased flexibility, processability and semitransparency. In OPVs, small organic molecules or polymers act as donor-acceptor pairs which absorb photons and create electrical charge. Our work address two primary challenges that are preventing the widespread adoption of OPVs relative to silicon-based solar cells: (1) inferior power conversion efficiency and (2) instability to sunlight. A "10-10" target has been long proposed that the OPVs need to reach >10% efficiency and 10

36th Conference on Neural Information Processing Systems (NeurIPS 2022).

years life time to be commercially viable [1]. Critical advances in non-fullerene acceptors[2] have provided new tandem organic solar cells with record-high efficiencies promising commercial viability (>19%)[3],[4]. However, the stability of OPVs has lagged far behind the "10-10" target. One recent work offers hope: the Forrest group reported one OPV system that remains stable nearly indefinitely when packaged under inert atmosphere, losing no efficiency after an extrapolated 9.3 years and losing only 20% efficiency after an extrapolated 27,000 years outdoors [5]. Nonetheless, there is not a single OPV system that meets the 10-10 target to date. It remains a critical challenge to design molecules that simultaneously optimize stability and efficiency, in part due to the lack of understanding how molecular structures dictate OPV degradation. In this proposal, we will address this grand challenge by developing an AI-guided autonomous materials design approach, wherein autonomous synthesis, autonomous property testing and machine learning form a close loop to drive molecular design towards the target properties. Scientifically, we aim to fill a key knowledge gap of how molecular structures determine OPV stability and efficiency. Technologically, we aim to meet the "10-10" target to make OPV a commercial reality for mitigating climate change.

2 Summary of prior works

Prior works on machine-learning-guided OPV discovery have relied heavily on incoherent literature data and simulated properties.[6, 7, 8] On the other hand, consistent, high quality and abundance of experimental data on molecular and device properties is crucial for machine learning based property prediction to ultimately enable meaningful inverse molecular design. Addressing this key bottleneck requires synthetic access to a wide array of candidate donor and acceptor structures in an automated, high throughput manner. Further, autonomous characterization of test OPV libraries is required at both the molecular and device levels. Autonomous synthesis and characterizations combined will generate sufficiently large and high quality datasets to enable machine learning-guided closed-loop materials discovery. Previously, Burke et. al.[9] have developed a series of automated small molecule synthesis instruments for building block-based synthesis of conjugated oligomers and small molecules. We seek to adapt this technology to the closed loop discovery of novel OPV materials.

We first sought to identify an initial target library of OPV molecules which we could use to probe key structure-property relationships as well as to provide initial datasets to kickstart AI property predictions. We based our initial molecular design on the latest design paradigm for hole transporting polymers comprised of a donor moiety and acceptor moiety linked by a pi conjugated bridge. A manual literature search of the most successful OPV chemistries was performed to identify donor, bridge, and acceptor moieties which were 1) high performing and 2) synthetically accessible using our building block based, iterative approach. We identified a synthetically accessible library of building blocks across our three categories. These blocks can be directly employed in our iterative synthesis machine to generate a library of unique donor-bridge-acceptor trimers to kickstart closed loop discovery. Combining these with commercially available blocks, we now have an accessible and modular building block set to enable the closed loop discovery of OPV molecules through building block-based automated synthesis.

In parallel, we have begun to design autonomous characterization of the photophysical properties of organic photovoltaic molecules.We have developed solution printing protocols for automatable fabrication of the optically active layer towards additive manufacturing of full OPV devices. This is necessary as the current practice of OPV fabrication and device testing is incompatible with automation, nor with large scale manufacturing. 3D printing provides an excellent programmable technology platform for addressing this challenge. Detailed accomplishments of our work in terms of Artificial Intelligence(AI), Synthesis, and characterization can be found in Appendix A.1-3.

3 Proposed research

One of the key challenges for creating a fully automated material discovery platform is the efficient synthesis and purification of ML-proposed candidate molecules that are poorly soluble. Previously the Synthesis team used a variety of strategies for synthesizing and purifying a diverse array of compounds covered by this project; however, solubility remains a challenge, and identifying a priori which method or solvent will be successful is an opportunity to better leverage ML. To overcome this challenge, the Synthesis and AI teams will use AI algorithms to predict synthetic and solubility parameters and to inform method/solvent selection in the Bayesian optimization scheme. To accomplish this goal, we

will identify quantifiable parameters that are advantageous for accelerating synthesis and purification. To predict molecular solubility in the initial small data limits (<100 molecules), we will use logistic regression and support vector machine models trained on our current molecular features because these models accurately predict categorical options given limited datasets. When larger datasets are obtained, we will switch to Bayesian regularized neural networks. In addition, we will take advantage of large, published solubility datasets and use transfer learning to improve prediction capability.

3.1 Artificial Intelligence

As synthetic methods improve and our ML models learn to accurately predict photostability, we will expand our exploration of OPV molecules to an extended chemical space to determine if the ML predictions derived from the initial chemical space apply to the full chemical space (Figure 1). This work will directly inform whether our approach might be a generalizable approach for future efforts to address inaccessibly large chemical spaces. To begin, we intentionally limited our chemical space to 22 donors and 100 acceptors (2200 oligomers) such that each monomer could be purchased on a finite budget and rapidly synthesized and learned over without shipping delays. However, the final optimal OPV of this approach is necessarily limited to the 2200 oligomers, even though the features learned are presumably transferable to the full chemical space of possibly millions of oligomers. To address this limitation, we will allow our ML algorithms to explore the full 22+ donor by 5000+ acceptor chemical space and sequentially purchase monomers that are predicted to form high-performing OPVs, expanding our limited dataset. Using this approach, we will assess the generalizability of the design rules learned by the ML algorithms, providing a new understanding to the OPV community. Moreover, addition of new donors and acceptors will non-linearly increase the size of our chemical space, leading to much broader learning. Initial work will focus on the best methods for selecting monomers to be added to our in-house monomer library by finding the optimal balance of exploring versus exploiting in the full chemical space using sample data sets.

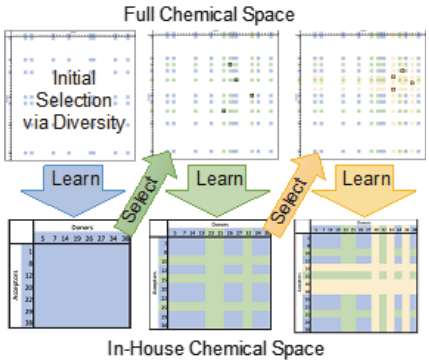

Figure 1: Expanding the ML-based predictive models to large chemical spaces

3.2 Synthesis

We first focused on generating the experimental and theoretical groundwork of an initial building block library of OPV moieties based on literature. A total of 22 donor-pi combinations and 100 diverse halides were prepared at a ready-to-use state, fully equipped for performing the automated iterative cross-coupling reactions. In this way, we explored a chemical space of 2200 trimer and dimer small molecules. We will continue to synthesize additional building blocks based on common OPV moieties, guided by AI-based predictions, while further expanding the synthetically accessible design space. Based on our prior work regarding solubility of OPV oligomers, we aim to expand the chemical space to include extended pi-bridges incorporating solubilizing side chains. We further aim to fully automate the purification process, while working towards interfacing the three independent units of closed-loop molecular discovery: AI-guided predictions, automated synthesis and purification, and solution and device characterization of optical properties. We aim to develop a fully autonomous closed-loop approach, which will be first-of-its-kind for OPVs (Figure2 b). To purify and characterize new molecules, we will interface the automated synthesis instruments with the automated preparatory LC/MS system. This step is crucial for analysis and purification of the crude reaction mixture by obtaining in-line mass spectrometry and UV-visible absorption data for initial chemical characterization of all target molecules. As this aspect of the process is currently achieved with the "human in the loop," we aim to remove the manual manipulations required in the purification process (Figure2 b).

We will also address molecular solubility, which we have learned is an important roadblock in terms of in purifying and isolating products. To address this challenge, we are using data mining of solution-phase data to develop a solubility database that can be integrated into the AI-guided algorithm. Adding solubility data into the closed-loop approach will improve the success rates of synthesized

molecules. We will identify and introduce solubility measurements into the database and use a final screening for solubility based on experimental results. Initial efforts will rely on evaluating solubility using a series of solvents and visually quantifying their solubility via solution turbidity. We further aim to automate this workflow, for example, building upon prior work using computer vision[10]. We will combine this advanced understanding of solubility with continued efforts to optimize general reaction conditions with maximized efficiency. Finally, we aim to extend the length (molecular weight) of oligomer to reach the "10-10" target, mainly because short oligomers are limited in terms of long-range charge transport and increased overlap with the solar spectrum. Here, we will perform iterative cross-coupling (ICC) using larger OPV building blocks (e.g. tetramers) containing solubilizing side chains. While developing this chemistry, we will compare the molecular properties of oligomers generated by ICC to those of a similar size generated using standard polymerization techniques, e.g. mesopolymers.

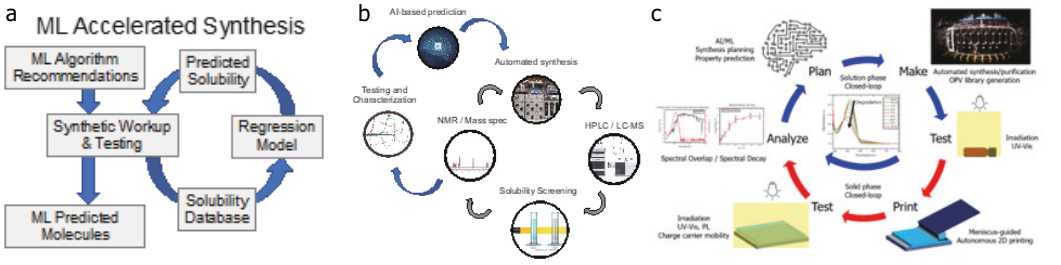

Figure 2: (a) Schematic workflow for AI-guided prediction of synthesis and solubility parameters. (b) Schematic of overall workflow integrating automated instrumentation: LC/MS for purification and solubility screening. (c) Schematic of autonomous fabrication and characterization platform consisting of 2 closed loops.

3.3 Automated characterization

We will continue building a database of solution-phase materials properties while further enabling OPV characterization in the solid phase. We will finish development of an in situ UV-Vis measurement setup, which will close the loop for AI-guided molecular design, synthesis, purification, and characterization. We will also build an autonomous characterization platform for film stability to augment the closed-loop process for solid-phase materials. In this way, the autonomous 2D printing platform will fabricate films, and the robotic arm utilized in the autonomous film printing platform will move the substrate to the automated imaging station to assess non-uniformity and defects by continuously capturing images. Uniform films will then be moved to an automated UV-Vis-NIR spectrometer, and the absorbance spectra of newly fabricated films will be determined. By utilizing the digital imaging and spectroscopic techniques, we aim to develop an analytical method to quantify the film thickness. UV-Vis measurements are not always capable of detecting trap sites and electrical defects in the film caused by the excitation of electrons that can undermine the charge carrier mobility of fabricated films thus the performance of OPV devices, so we will add automated photoluminescence detectors and a 4-point probe mobility testing station. Following analysis, the films will be transferred to a custom-built environmental chamber capable of humidity control and oxygen level tracking. This environmental chamber simulates the encapsulated environment to minimize the effect of humidity and oxygen to degradation and to analyze the intrinsic stability of printed films. Films in the environmental chamber will be exposed to AM 1.5G solar radiation emitted by automated 1 Sun solar simulator to track degradation. Degraded films will be transferred to the testing stations on an hourly basis to check their properties over time until they reach T80. The design of this platform will be modular for each compartment and controlled through a python-based custom program to integrate the platform with other workflows in the closed loop including autonomous synthesis, device fabrication, and characterization.

4 Conclusion

We aim to develop a fully autonomous closed-loop framework for OPV discovery, which will be the first-of-its-kind in the field to realize AI-guided design of OPV molecules that meet the "10-10" target.

This will require integrating all three major components of the proposal: autonomous molecular design and synthesis, autonomous molecular and device property testing, and interpretable AI for high fidelity property prediction and discovery of molecular design rules. In the long term, we will establish a living database of electronically active organic molecules with measured and predicted properties that can be interfaced with AI-based property prediction. In parallel, we will fully automate the material synthesis, device manufacture and testing to map functional properties to molecular and building block properties to ultimately realize inverse materials design for targeted function such as high-power conversion efficiency and device stability. With our autonomous materials discovery platform for organic photovoltaics, it would be able to identify the structure-property relationship of electronically active organic molecules, enabling the autonomous system to ultimately inverse-design the optimal organic photovoltaic material with intrinsic stability meeting the "10-10" target.

References

[1] Roar Søndergaard, Markus Hösel, Dechan Angmo, Thue T. Larsen-Olsen, and Frederik C. Krebs. Roll-to-roll fabrication of polymer solar cells. *Materials Today*, 15(1):36–49, 2012.

[2] Jianhui Hou, Olle Inganäs, Richard H Friend, and Feng Gao. Organic solar cells based on non-fullerene acceptors. *Nature materials*, 17(2):119–128, 2018.

[3] Lingxian Meng, Yamin Zhang, Xiangjian Wan, Chenxi Li, Xin Zhang, Yanbo Wang, Xin Ke, Zuo Xiao, Liming Ding, Ruoxi Xia, et al. Organic and solution-processed tandem solar cells with 17.3% efficiency. *Science*, 361(6407):1094–1098, 2018.

[4] Lei Zhu, Ming Zhang, Jinqiu Xu, Chao Li, Jun Yan, Guanqing Zhou, Wenkai Zhong, Tianyu Hao, Jiali Song, Xiaonan Xue, et al. Single-junction organic solar cells with over 19% efficiency enabled by a refined double-fibril network morphology. *Nature Materials*, 21(6):656–663, 2022.

[5] Quinn Burlingame, Xiaheng Huang, Xiao Liu, Changyeong Jeong, Caleb Coburn, and Stephen R Forrest. Intrinsically stable organic solar cells under high-intensity illumination. *Nature*, 573(7774):394–397, 2019.

[6] Johannes Hachmann, Roberto Olivares-Amaya, Sule Atahan-Evrenk, Carlos Amador-Bedolla, Roel S Sánchez-Carrera, Aryeh Gold-Parker, Leslie Vogt, Anna M Brockway, and Alán Aspuru-Guzik. The harvard clean energy project: large-scale computational screening and design of organic photovoltaics on the world community grid. *The Journal of Physical Chemistry Letters*, 2(17):2241–2251, 2011.

[7] Steven A Lopez, Benjamin Sanchez-Lengeling, Julio de Goes Soares, and Alán Aspuru-Guzik. Design principles and top non-fullerene acceptor candidates for organic photovoltaics. *Joule*, 1(4):857–870, 2017.

[8] Wenbo Sun, Yujie Zheng, Ke Yang, Qi Zhang, Akeel A Shah, Zhou Wu, Yuyang Sun, Liang Feng, Dongyang Chen, Zeyun Xiao, et al. Machine learning–assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Science advances*, 5(11):eaay4275, 2019.

[9] Junqi Li, Steven G Ballmer, Eric P Gillis, Seiko Fujii, Michael J Schmidt, Andrea ME Palazzolo, Jonathan W Lehmann, Greg F Morehouse, and Martin D Burke. Synthesis of many different types of organic small molecules using one automated process. *Science*, 347(6227):1221–1226, 2015.

[10] Parisa Shiri, Veronica Lai, Tara Zepel, Daniel Griffin, Jonathan Reifman, Sean Clark, Shad Grunert, Lars PE Yunker, Sebastian Steiner, Henry Situ, et al. Automated solubility screening platform using computer vision. *Iscience*, 24(3):102176, 2021.

[11] Connor P Delaney, Vincent M Kassel, and Scott E Denmark. Potassium trimethylsilanolate enables rapid, homogeneous suzuki- miyaura cross-coupling of boronic esters. *ACS catalysis*, 10(1):73–80, 2019.

[12] Paul A Cox, Marc Reid, Andrew G Leach, Andrew D Campbell, Edward J King, and Guy C Lloyd-Jones. Base-catalyzed aryl-b (oh) 2 protodeboronation revisited: from concerted proton transfer to liberation of a transient aryl anion. *Journal of the American Chemical Society*, 139(37):13156–13165, 2017.

A Appendix

In this section we present our previous accomplishments towards building an autonomous material discovery platform for organic solar cells in terms of AI, synthesis, and characterization aspect.

A.1 Artificial Intelligence (AI)

We adapted an established machine learning (ML) algorithm to discover new OPV oligomers with high photostability and power conversion efficiency (PCE) (figure 3, a). By utilizing a Bayesian optimization framework in which physicochemical descriptors of OPV candidates, we have guided a search through a large molecular combinatorial space while maintaining a customizable tradeoff between exploitative and explorative sampling (figure 3, b). A major computational hurdle was encountered by the need to generate molecular descriptors for all candidates in the proposed chemical search space. Initial estimates indicated that density functional theory (DFT) calculations necessary to produce electronic descriptors for 100k oligomers would require 300k CPU-days. To circumvent this computational bottleneck, we computed DFT descriptors for only the individual 22 donors and 5000+ acceptor blocks and used the concatenated features of each block to featurize corresponding oligomers. Following featurization of the chemical space, we downselected 100 acceptors from the initial library of 5000+, while preserving as much chemical diversity as possible. Although many of the molecules were successfully made on the synthesis robots, some proved to be undeliverable with the first-gen synthesis protocol. This created a computational challenge because ML is designed to recommend a set of molecules that balances exploration and exploitation of chemical space. To compensate the failure to access all of the selected molecules for each round, we adapted our ML algorithm to recommend several 'back-up' molecules (nearest neighbors in feature space) for each recommended OPV, which provided the synthesis team with alternate target molecules that maintain a balanced portfolio of exploration and exploitation and maximize the rate of discovery of high-performing OPVs.

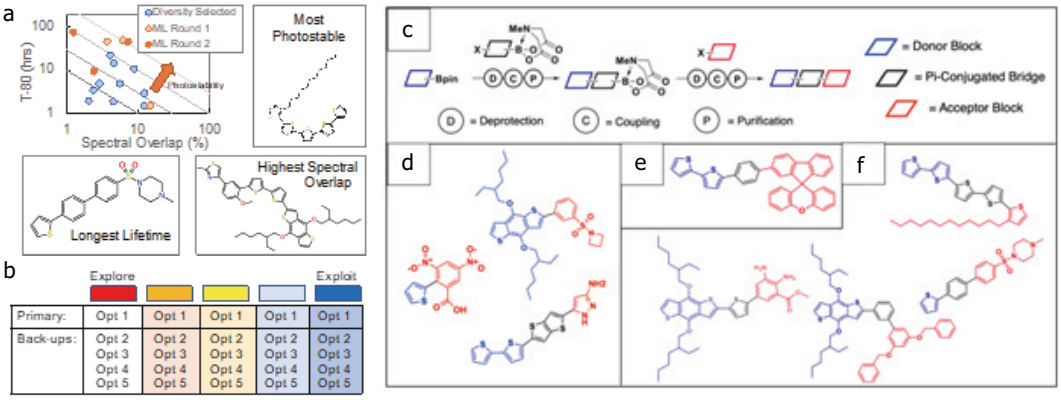

Figure 3: (a-b) Bayesian optimization framework to guide OPV discovery based on photo-physical properties (lifetime, spectral overlap (c) Iterative cross coupling reaction scheme. (d-f) Sampling of successfully synthesized molecules in reaction conditions: (d) 5 mol% Pd X Phos, 7.5 equiv Na2CO3, 5:1 dioxane:H2O, 100C, 12 h; (e) 3 equiv. TMSOK, 10 mol% Pd (P(tBu3)3) G3, THF, 25C, 2 h; (f) same conditions as (e) but at 60C, for 1 h and additional 1 equiv. pinacol.

A.2 Automated synthesis

Subsets of structurally complex and diverse OPV molecules were synthesized in batches of 8-12 molecules suggested by the AI sub-team and sent to the characterization sub-team for analysis of solution-phase optical properties. A generalized set of reaction conditions and a custom-built automated synthesizer as Burke et. al. has developed[9] were used to increase the success rate of molecular synthesis in each round. Using reaction conditions previously discovered by our collaborators, we successfully isolated and delivered an initial 13 out of 30 (43%) of the molecules suggested by the AI team. However, we encountered an important challenge because some of the reaction yields were relatively low for this complex and diverse chemical space due to the relative insolubility of intermediates and/or instability of the boronic acids generated during the reaction (Figures 3 c,d). Additional challenges were encountered in purifying and isolating the reaction products due to the broad chemical diversity and low solubility of some target molecules. To overcome these challenges, synthesis conditions were optimized by implementing recently discovered rapid Suzuki reaction conditions utilizing potassium trimethylsilanolate (TMSOK) as a base[11],

where the reaction conditions are anhydrous and homogenous, which greatly limits the mechanisms of protodeboronation.[12] We found that the protected boronates utilized in the automated iterative synthesis are rapidly transligated to reactive but relatively stable pinacol boronic esters to afford this rapid, homogenous reactivity. After optimization of this process, the general reaction conditions were modified including a decrease in temperature from 100°C to 25-60'C with a >10-fold decrease in reaction time (1-2 h) (Figures 2c,d). Application of these reaction conditions to prior molecules suggested by the AI team that could not be previously synthesized using the prior conditions has so far resulted in the delivery of 18 out of 36 (50%) of the suggested molecules. In parallel, to develop a generalized purification methods for suggested molecule, we directly interfaced the automated synthesis hardware with the automated liquid chromatography/mass spectrometry (LC/MS) system, thereby enabling analysis and purification of crude reaction mixtures while also obtaining chemical characterization (MS) and UV-Vis absorption data. However, we observed that crude reaction mixtures in this chemical space do not uniformly ionize in electrospray ionization (ESI) MS, limiting the prospects of using LC/MS as a frontline analytical readout to inform the purification. Instead, we used in-line filtration methods (silica gel) to automate the work-up and to directly interface synthesis with automated purification apparatus.

A.3 Automated characterization

We established a workflow for autonomous characterization of the photophysical properties of OPVs. We defined an objective function for AI-based optimization as spectral overlap (SO) × T80 for synthesized molecules, where T80 is the time required for the molecular absorbance spectrum to decay to 80% of the initial spectrum. Suitable OPV molecules should efficiently absorb sunlight while maintaining electrochemical properties, so we determined the UV-Vis absorbance spectra of the molecules in the solution phase with the AM 1.5G solar spectrum. By analyzing the decay of the absorbance spectra over time, T80 values were also determined. Using this approach, we analyzed SO and T80 of OPV molecules synthesized by the Synthesis team in the solution phase and passed the data to the AI team to learn from and enable next-round predictions (Figures 4 a-c). In addition to photophysical property characterization, we also developed autonomous methods for solution-phase printing of solar cells. A complete OPV device consists of a stack of multilayered films. Because the properties and degradation behavior of OPV molecules are different in solution and film phases, we developed an autonomous film fabrication process (Figure 4 d) to enable rapid characterization in the solid phase. By incorporating a robotic arm, a stir plate, syringe pumps and custom-made parts including a substrate handler and a liquid dispenser, we can reliably handle substrates between different stations, dilute precursor solutions, and dispense the solution on top of the substrate prior to film deposition. In this way, we developed an automated meniscus-guided 2D printer by utilizing linear stages and some custom-made appliances including an ultrasonic cleaner to automate film printing and blade cleaning process. Using the autonomous 2D film printing platform, we are now able to fabricate the optically active semiconducting layer for film-phase degradation experiments of synthesized molecules. Using an AI-guided closed-loop approach, it will be possible to correlate the photostability and optoelectronic properties of OPVs in the solution and the solid phases and to utilize the solution-phase experimental data as a fingerprint for solid-phase properties.

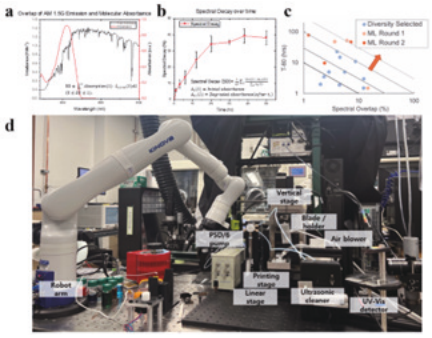

Figure 4: (a) Analysis of spectral overlap based on absorbance spectra of synthesized molecule in the solution phase. (b) Analysis of spectral decay based on UV-Vis absorbance spectra of synthesized molecule in the solution phase. (c) Mapped space of synthesized molecules characterized by spectral overlap and T80 in the solution phase. (d) Custom-built autonomous 2D printing platform