Supporting Information

Screening interface passivation materials intelligently through machine learning for highly efficient perovskite solar cells

Wu Liu,[a,b] Yao Lu,[a,b] Dong Wei,[c], Xiaomin Huo, [a,b] Xiaofeng Huang, [a,b] Yaoyao Li, [a,b] Juan Meng, [a,b] Suling Zhao, [a,b] Bo Qiao, [a,b] Zhiqin Liang, [a,b] Zheng Xu, [a,b],* Dandan Song [a,b],**

[a] Key Laboratory of Luminescence and Optical Information, Beijing Jiaotong University, Ministry of Education, Beijing 100044, China

[b] Institute of Optoelectronics Technology, Beijing Jiaotong University, Beijing 100044, China

[c] College of Physics and Energy, Fujian Normal University, Fuzhou, 350117, China.

E-mail: ddsong@bjtu.edu.cn (Dandan Song); zhengxu@bjtu.edu.cn (Zheng Xu); q397983012@126.com (Dong Wei)

1. Experimental

Materials Preparation: PbI_2 (purity > 99.9985%), MAI (purity > 99.995%), FAI (purity > 99.995%), MABr (purity > 99.995%), and spiro-OMeTAD (purity > 99.5%) were purchased from Liaoning Youxuan Corp., Diethyl ether (DEE), chlorobenzene (CB), and lithium bis (trifluoromethylsulfonyl) imide salt (Li-TFSI) were purchased from Sigma-Aldrich. Dimethylformamide (DMF, purity > 99.8%) and dimethyl sulfoxide (DMSO, purity > 99.7%) were purchased from Aladdin reagent. All Materials are used as-received.

Perovskite and Spiro-OMeTAD Solution: Perovskite precursor solution of $FA_{0.6}MA_{0.4}PbI_3$ was composed by dissolving FAI (0.6 mmol), MABr (0.07 mmol), MAI (0.33 mmol), and PbI_2 (1.07 mmol) in 1 mL mixed solvent of DMF and DMSO (volume ratio of 4:1). The spiro-OMeTAD solution was composed of spiro-OMeTAD and Li-TFSI. The solutions mentioned above were stirred at 25 °C for 12 h and then filtered through a 0.45 μm syringe filter.

Fabrication of Solar Cells: The pre-cleaned FTO substrates were dried with nitrogen and treated with UV ozone for 10 min. The TiO$_2$ layer was prepared by the water bath deposition as the ETL. The perovskite precursor solution was spin-coated at 7000 *rpm* for 25 s on the FTO/TiO$_2$ substrates with diethyl ether (DEE) as the anti-solvent. After the spin coating, the perovskite films were annealed at 150 °C for 10 min and then cooled to room temperature. The films were cooled down to room temperature for the surface passivation treatment and treated with MAI or PEAI solution. The Spiro-OMeTAD solution was coated on perovskite films at 4000 *rpm* for 30 s. Finally, the devices were completed with a 70 nm thick gold counter electrode using thermal evaporation.

Extraction of parameters from J-V curves and SCLC measurement: Planar structured PSCs can be treated as a single heterojunction diode. The electric parameters of the PSCs, including the ideality factor of the diode (*m*), the series resistance (R_S), and the reverse saturation current (J_0) of the PSCs, can be calculated according to the diode equation.[1,2]

$$-\frac{dV}{dJ} = \frac{mk_BT}{e}(J_{SC} - J)^{-1} + R_S \qquad (1)$$

$$\ln(J_{SC} - J) = \frac{e}{mk_BT}(V + R_S \times J) + \ln(J_0) \qquad (2)$$

where *J* is the current density flow through the external load, *e* is the elementary charge, and K_B and *T* are the Boltzmann constant and the absolute temperature, respectively.

The trap state density N_{trap} can be determined from the trap-filled limit voltage (V_{TFL}) by Equation (3):

$$N_{trap} = \frac{2\varepsilon_0\varepsilon}{qL^2}V_{TFL} \qquad (3)$$

where *q* is the electron charge, N_{trap} is the trap state density, *L* is the thickness of the perovskite layer, ε is the relative dielectric constant for FA$_{0.6}$MA$_{0.4}$PbI$_3$, and ε_0 is the vacuum permittivity. As we can see, the only variable in the equation is V_{TFL}, which can be extracted from the *J-V* plots in the SCLC region.

2. DFT calculation

The structural optimization and electronic structure calculations were carried out by Cambridge Serial Total Energy Package[3,4] (CASTEP) in *Materials studio*. The generalized gradient approximation (GGA) of the Perdew-Burke-Ernzerhof (PBE) functional was employed.[5] Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was used for the structural optimization of the model with the following optimization parameters: the calculation was expanded by using the ultrasoft pseudopotential with a cutoff energy of 435 eV, and the total energy was converged to 2×10^{-5} eV. The structural optimization was optimized until the force tolerance on each atom was smaller than $0.05 \mathrm{eV\AA^{-1}}$, the stress tolerance was smaller than 0.1 GPa, and the displacement tolerance was smaller than 0.002 Å. The Monkhorst-Pack grids with the actual spacing of 0.041 $\mathrm{\AA^{-1}}$ and SCF tolerance of $2\times e^{-6}$ eV/atom were used in all DFT simulations.

$FA_{0.6}MA_{0.4}PbI_3$ possesses a cubic structure, with the space group Pm-3m at room temperature. A 3×3×1 supper cell and a 20 Å vacuum slab were employed to investigate the adsorption of different additives. All additives were placed on the supper cell surface to optimize to convergence.

A large number of studies have shown that high-performance PSCs often adopt the strategy of excessive PbI_2. At the same time, according to the results of previous DFT studies, the PbX_2 terminal on the 001 surfaces has the lowest energy among all types of surfaces.[6] In the case of using PEAI as the interface modification material, PEAI adsorbs on the $[PbI_4]^{2-}$ surface of the perovskite during interface passivation. Meanwhile, since the Pb-I bond is stronger than the interaction between PEA-I, we consider the weakest binding energy of the PEA-I bonds as the final binding energy value to compare different interface passivation materials. Therefore, we use the I-terminal as the surface state, and PEA^+ is placed on the perovskite surface.[7] Binding energies were calculated based on the surface relaxation approach, following

$$E_{binding\ energy} = E_{all} - E_{surface} - E_{additive}$$

Where $E_{\text{binding energy}}$ is the final binding energy between additives and perovskite surface, E_{all} is the total energy of the system after relaxation of additives adsorbed on perovskite surface, E_{surface} is the surface energy of the perovskite without additives, E_{additive} is the energy of the additive. The settings in the cases of using other organic salts are similar to these of PEAI.

3. Machine learning

Dataset Preparation: For dataset preparation, we mainly focus on utilizing interface materials at the perovskite/Spiro-OMeTAD HTL interface. The data of the ML model were collected from the published literature, including various publishing groups such as Wiley, Elsevier, RSC, Nature, and Science. To ensure the uniformity of data collection standards and to establish a high-quality dataset, we removed inorganic interface material to avoid its noise to organic functional groups. Only the reports within the last five years were kept to reflect the current development of interface engineering. In addition, devices with PCEs above 18% were selected to reduce data dispersion while ensuring the learning accuracy of the high-PCE mapping models. To minimize the effect of the standard device and clearly show the role of the interface material, we also use the control device performance (the control PCE) as the input feature. The ion ratios in the perovskite precursors were counted rather than those suggested in the films. The data points using the two-step film formation method also were washed away.

Feature Engineering: Several factors determine a material's or device's properties, also called features. Feature extraction follows three principles: highly correlated to the output, easy to obtain, and minimal in number. The SHAP value analysis and feature importance analysis from the algorithm model were used to reflect the influence of the feature of each sample, and the correlation analysis in statistical analysis was used to reflect the degree of linear correlation between the two variables. We ranked the initial more than 300 features through the SHAP value, feature importance, and correlation analysis. We screened the most critical 15 features from more than 300 features describing perovskite, interface material, and control device

performance. Therefore, we constructed a set of 15-dimensional features as the input of the photovoltaic parameter prediction model (Table 1), which reduces the overfitting problem caused by the high complexity of the model.

Machine learning settings: *Rstudio* was employed as the platform for machine learning. Linear regression (LR), neural network (NN), random forest (RF), and extreme gradient boosting (XGBoost) algorithms were used based on *glm*, *neuralnet*, *randomForest,* and *xgboost* functions, respectively. The performance of the algorithms was evaluated by the root mean square error (RMSE) and Pearson's coefficient (*r*-value) on the test set. Here,

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(X_i - Y_i)^2}{n}} \qquad (4)$$

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}} \qquad (5)$$

X_i, Y_i, $\overline{X}$, $\overline{Y}$, and n represent the ith value of the experimental dataset, the ith value of the predicted dataset, the mean value of the experimental dataset, the mean value of the predicted dataset, and the number of the dataset points, respectively. The test set accounted for 30% of the entire database and was mutually exclusive with the training set. We use 10-fold cross-validation to optimize the hyperparameters, dividing the train set into ten parts (90% data points for training and 10% for validation) and learning ten times. The model performing the lowest average RMSE on the validation sets was screened for testing on the test set and further use. In detail, the NN model had four hidden layers, which had 100, 50, 20, and 2 neurons, respectively; the tree number in the RF model was 5000. The max depth and the number of rounds in the XGBoost model were 10 and 25, respectively.
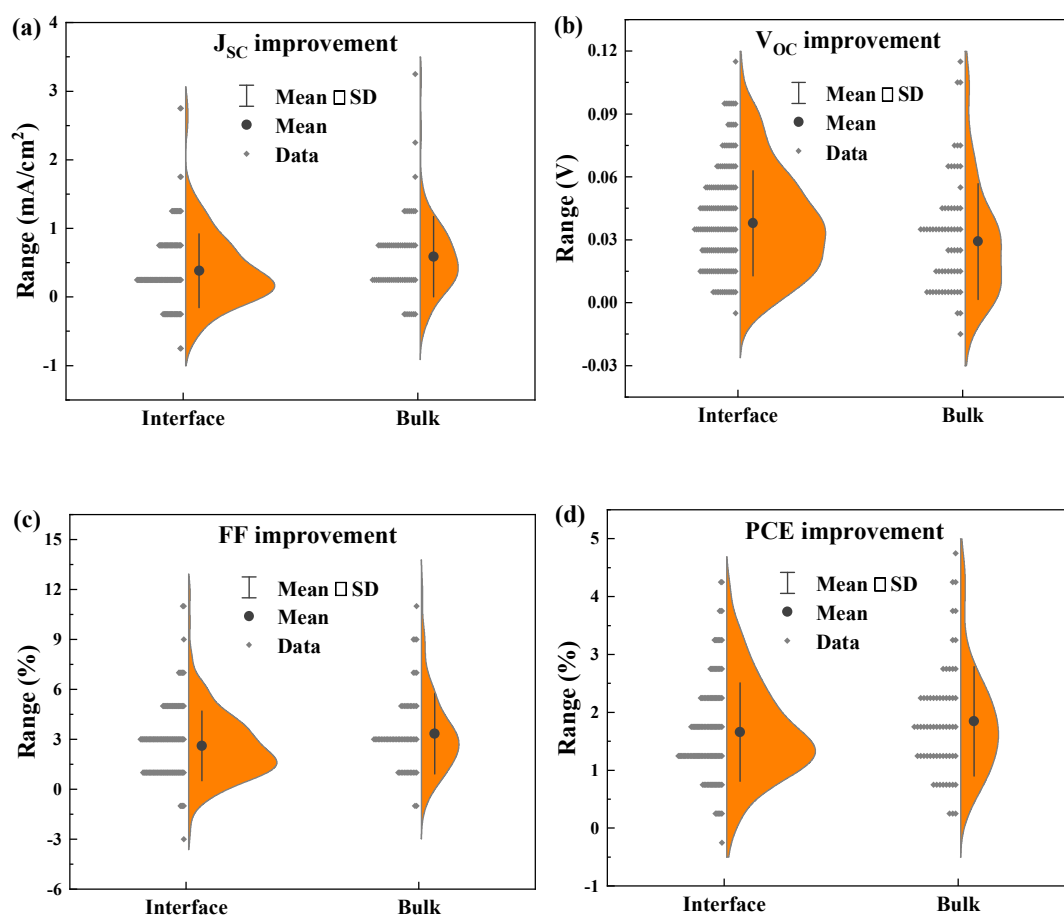
4. Supplementary results

Figure S1. Performance improvements of different passivation strategies based on more than 200 highly efficient *nip*-type PSCs. Here, the device performance is generally referred to as the values obtained in the reverse scan, except the scan direction is not clearly stated. The performance improvements were the differences between the modified and control devices. The interface approach refers to modifying the perovskite film before or after the film fabrication of the perovskite layer. In contrast, the bulk approach refers to adding additives into the perovskite precursor solution or during the film fabrication of the perovskite layer.

Table S1. Reports on the interface modification at the perovskite/Spiro-OMeTAD interface of the PSCs

ID	Device PCE	Control device PCE	Interface material	Perovskite precursor solution			References' DOI
				A site cation and ratio	A2X	P2X	
1	20.50	19.20	S	$Cs_{0.1}FA_{0.9}$	1:3	1.05:3	10.1038/s41467-018-06709-w
2	20.50	19.20	N	$Cs_{0.1}FA_{0.9}$	1:3	1.05:3	10.1038/s41467-018-06709-w
3	20.90	19.20	SN	$Cs_{0.1}FA_{0.9}$	1:3	1.05:3	10.1038/s41467-018-06709-w
4	21.01	19.45	PEIm	$Cs_{0.1016}FA_{0.7813}MA_{0.1172}$	0.965:3	1.018:3	10.1039/D0TA05496H
5	19.01	15.01	P3HT	MA	1.8:3	0.6:3	10.1016/j.jpowsour.2017.12.082
6	18.61	15.01	PTAA	MA	1.8:3	0.6:3	10.1016/j.jpowsour.2017.12.082
7	18.35	15.01	MEH-PPV	MA	1.8:3	0.6:3	10.1016/j.jpowsour.2017.12.082
8	17.83	15.01	Poly-TPD	MA	1.8:3	0.6:3	10.1016/j.jpowsour.2017.12.082
9	17.66	15.01	PBDTTT-CT	MA	1.8:3	0.6:3	10.1016/j.jpowsour.2017.12.082
10	19.57	17.48	MMI	MA	1:3	1:3	10.1002/adma.201800544
11	19.20	15.70	Benzylamine	FA	1:3	1:3	10.1002/adma.201603062
12	18.10	16.10	PMMA	MA	1:3	1:3	10.1021/acs.jpcc.6b12137
13	18.87	18.08	TP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
14	18.82	18.08	MTP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
15	19.17	18.08	ETP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
16	19.44	18.08	BTP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
17	19.89	18.08	HTP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
18	19.20	18.08	DTP	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/aenm.201703143
19	20.50	18.90	FABr	$FA_{0.8462}MA_{0.1538}$	0.975:3	1.013:3	10.1039/C6EE03182J

20	18.51	17.02	PEAI	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.989:3	1.005:3	10.1002/adfm.201706923
21	22.00	20.70	C$_4$Br	$FA_{0.6965}MA_{0.3035}$	1.14:3	0.93:3	10.1039/C9EE00751B
22	22.40	20.70	C$_6$Br	$FA_{0.6965}MA_{0.3035}$	1.14:3	0.93:3	10.1039/C9EE00751B
23	22.10	20.70	C$_8$Br	$FA_{0.6965}MA_{0.3035}$	1.14:3	0.93:3	10.1039/C9EE00751B
24	20.28	18.35	2-MP	MA	0.968:3	1.016:3	10.1002/aenm.201803573
25	18.74	18.35	PTT	MA	0.968:3	1.016:3	10.1002/aenm.201803573
26	19.02	18.35	Py	MA	0.968:3	1.016:3	10.1002/aenm.201803573
27	22.30	20.72	EAI	$Cs_{0.07}FA_{0.9}MA_{0.03}$	0.98:3	1.01:3	10.1038/s41467-019-10985-5
28	21.60	20.72	IAI	$Cs_{0.07}FA_{0.9}MA_{0.03}$	0.98:3	1.01:3	10.1038/s41467-019-10985-5
29	20.90	20.72	GuaI	$Cs_{0.07}FA_{0.9}MA_{0.03}$	0.98:3	1.01:3	10.1038/s41467-019-10985-5
30	20.54	19.22	FPEAI	$Cs_{0.0617}FA_{0.7819}MA_{0.1564}$	0.964:3	1.018:3	10.1002/aenm.201802595
31	19.20	16.40	tBP	$FA_{0.85}MA_{0.15}$	1:3	1:3	10.1039/C9EE01773A
32	19.43	18.83	BAI	$Cs_{0.0477}FA_{0.8095}MA_{0.1428}$	0.975:3	1.012:3	10.1021/acsami.9b17930
33	20.62	18.83	HAI	$Cs_{0.0477}FA_{0.8095}MA_{0.1428}$	0.975:3	1.012:3	10.1021/acsami.9b17930
34	21.19	19.52	EPC	$FA_{0.7159}MA_{0.2841}$	1.2:3	0.9:3	10.1039/D0TA02222E
35	20.50	20.45	TFDIB	$Cs_{0.05}FA_{0.855}MA_{0.095}$	0.971:3	1:3	10.1021/jacs.9b13701
36	19.60	18.50	HS-Ph-CN	$Cs_{0.0514}FA_{0.7905}MA_{0.1581}$	0.972:3	1.014:3	10.1039/C8EE00754C
37	20.00	19.00	HS-Ph-NO$_2$	$Cs_{0.0514}FA_{0.7905}MA_{0.1581}$	0.972:3	1.014:3	10.1039/C8EE00754C
38	19.10	19.00	HS-Ph-SCH$_3$	$Cs_{0.0514}FA_{0.7905}MA_{0.1581}$	0.972:3	1.014:3	10.1039/C8EE00754C
39	19.40	19.00	HS-Ph-OCH$_3$	$Cs_{0.0514}FA_{0.7905}MA_{0.1581}$	0.972:3	1.014:3	10.1039/C8EE00754C
40	20.62	17.43	PHI	$Cs_{0.05}FA_{0.7885}MA_{0.1615}$	0.976:3	1.012:3	10.1021/acsami.0c10448
41	20.80	20.28	PMMA	$Cs_{0.0636}Rb_{0.0273}FA_{0.7692}MA_{0.1399}$	1.014:3	0.993:3	10.1002/aenm.201801208

42	21.09	19.55	ClTPPPF$_6$	Cs$_{0.0367}$Rb$_{0.015}$FA$_{0.6977}$MA$_{0.2506}$	1.201:3	0.9:3	10.1002/eom2.12158
43	22.09	19.55	BrTPPPF$_6$	Cs$_{0.0367}$Rb$_{0.015}$FA$_{0.6977}$MA$_{0.2506}$	1.201:3	0.9:3	10.1002/eom2.12158
44	22.48	19.93	CHAI	FA	1.304:3	0.848:3	10.1002/aenm.202102236
45	23.07	19.93	CHMAI	FA	1.304:3	0.848:3	10.1002/aenm.202102236
46	21.70	20.10	(HAD)I$_2$	Cs$_{0.0502}$FA$_{0.8094}$MA$_{0.1405}$	0.945:3	1.027:3	10.1002/aenm.202102973
47	22.60	20.10	(EDBE)I$_2$	Cs$_{0.0502}$FA$_{0.8094}$MA$_{0.1405}$	0.945:3	1.027:3	10.1002/aenm.202102973
48	20.56	18.14	CDCA	Cs$_{0.0625}$FA$_{0.7813}$MA$_{0.1563}$	0.99:3	1.005:3	10.1016/j.jpowsour.2020.228502
49	21.94	19.17	C8-BTBT	Cs$_{0.05}$FA$_{0.85}$MA$_{0.1}$	0.981:3	1.009:3	10.1021/acsenergylett.1c01898
50	21.29	19.71	MTDAA	Cs$_{0.0367}$Rb$_{0.015}$FA$_{0.6977}$MA$_{0.2506}$	1.201:3	0.9:3	10.1021/acsenergylett.1c00794
51	21.20	20.52	ADAHCl	Cs$_{0.0601}$FA$_{0.7991}$MA$_{0.1408}$	1.015:3	0.993:3	10.1002/aenm.201803587
52	20.46	19.24	Polystyrene	Cs$_{0.0588}$FA$_{0.7843}$MA$_{0.1569}$	0.977:3	1.011:3	10.1021/acsami.8b04776
53	18.95	16.99	CTABr	MA	1:3	1:3	10.1039/C9TA02631B
54	20.28	19.62	CTABr	FA$_{0.95}$MA$_{0.05}$	1:3	1:3	10.1039/C9TA02631B
55	20.13	18.52	TCPBr	MA	1:3	1:3	10.1039/C9TA12597C
56	19.41	18.52	TCPI	MA	1:3	1:3	10.1039/C9TA12597C
57	20.40	16.80	PMMA	MA	1.001:3	0.999:3	10.1002/admi.201701256
58	21.90	20.90	PTPD	Cs$_{0.0564}$FA$_{0.9}$MA$_{0.05}$	1:3	1:3	10.1002/adma.201807435
59	21.60	19.30	Tetracene	Cs$_{0.06}$FA$_{0.7935}$MA$_{0.1501}$	1.013:3	1.021:3	10.1126/sciadv.aav2012
60	20.47	19.01	PABA	MA	1:3	1:3	10.1002/admi.201901584
61	20.93	19.32	ADA	Cs$_{0.0497}$FA$_{0.8077}$MA$_{0.1425}$	1.006:3	0.997:3	10.1002/aenm.201800275
62	20.47	19.32	AD	Cs$_{0.0497}$FA$_{0.8077}$MA$_{0.1425}$	1.006:3	0.997:3	10.1002/aenm.201800275
63	20.90	19.80	PVP	Cs$_{0.04}$FA$_{0.8}$MA$_{0.16}$	0.964:3	1.018:3	10.1021/acsomega.8b00555

64	21.11	19.61	B_2Cat_2	$Cs_{0.0465}FA_{0.8450}MA_{0.1085}$	0.97:3	1.015:3	10.1002/adma.201805085
65	21.53	19.48	OLA	$Cs_{0.0378}FA_{0.6280}MA_{0.3342}$	1.194:3	0.903:3	10.1021/acsenergylett.0c00279
66	20.04	19.10	$BMIMBF_4$	$Cs_{0.0491}FA_{0.7953}MA_{0.1556}$	0.995:3	1.003:3	10.1016/j.orgel.2020.105805
67	19.17	18.26	$POSS-NH_2$	$FA_{0.8333}MA_{0.1667}$	0.947:3	1.026:3	10.1021/acsaem.9b00050
68	18.38	18.26	POSS-SH	$FA_{0.8333}MA_{0.1667}$	0.947:3	1.026:3	10.1021/acsaem.9b00050
69	19.02	17.20	BEDCE	MA	1.005:3	0.997:3	10.1039/C8TA09724K
70	20.90	19.70	cesium acetate	$Cs_{0.05}FA_{0.8}MA_{0.15}$	1:3	1:3	10.1021/acsami.8b10616
71	18.31	17.65	Cs-oleate	$Cs_{0.0476}FA_{0.7905}MA_{0.1619}$	1.033:3	0.984:3	10.1021/acsami.9b08026
72	18.97	17.92	TAI	$FA_{0.9}MA_{0.1}$	0.938:3	1.031:3	10.1021/acsenergylett.9b00930
73	19.89	18.61	ImI	MA	1:3	1:3	10.1016/j.nanoen.2018.05.035
74	19.10	17.20	P(VDF-TrFE)	$Cs_{0.2}FA_{0.8}$	1:3	1:3	10.1246/cl.190692
75	18.83	17.51	PD-10-DTTE-7	MA	1:3	1:3	10.1002/solr.201800232
76	20.56	19.25	TFMBA	$Cs_{0.0497}FA_{0.8077}MA_{0.1425}$	0.985:3	1.008:3	10.1016/j.cej.2020.126712
77	23.25	21.09	$DMIMPF_6$	$Cs_{0.08}FA_{0.92}$	1:3	1:3	10.1002/anie.202010987
78	20.16	17.94	PEACl	$FA_{0.1}MA_{0.9}$	1.286:3	0.857:3	10.1021/acsaem.1c02210
79	20.70	19.30	BAI	MA	1.026:3	0.987:3	10.1039/D1RA02260A
80	19.42	18.41	SDBS	MA	1:3	1:3	10.1021/acsami.0c14732
81	20.83	19.45	PyNa+	$Cs_{0.0517}FA_{0.8621}MA_{0.0862}$	0.978:3	1.011:3	10.1002/solr.202100416
82	22.66	20.61	OAI	$Cs_{0.0636}Rb_{0.0273}FA_{0.769}MA_{0.14}$	1.014:3	0.993:3	10.1002/adfm.202104251
83	23.38	20.61	OABr	$Cs_{0.0636}Rb_{0.0273}FA_{0.769}MA_{0.14}$	1.014:3	0.993:3	10.1002/adfm.202104251
84	23.62	20.61	OACl	$Cs_{0.064}Rb_{0.0273}FA_{0.769}MA_{0.14}$	1.014:3	0.993:3	10.1002/adfm.202104251
85	21.95	19.76	CBAH	FA	1.124:3	0.938:3	10.1002/smll.202104100

|							
86	24.33	22.67	[EMIM]Br	FA	0.959:3	1.021:3	10.1002/aenm.202103491
87	20.15	18.64	DMEDAI$_2$	MA	1:3	1:3	10.1021/acsami.9b17851
88	21.37	19.36	Poly TPD	FA$_{0.85}$MA$_{0.15}$	1:3	1:3	10.1002/adma.202006087
89	17.49	14.16	Poly TPD	FA	1:3	1:3	10.1002/adma.202006087
90	22.16	20.62	FEAI	Cs$_{0.04}$FA$_{0.9201}$MA$_{0.0398}$	0.958:3	1.021:3	10.1126/sciadv.aaw2543
91	19.10	18.05	TPPO	Cs$_{0.05}$FA$_{0.8}$MA$_{0.15}$	1:3	1:3	10.1002/adfm.201910710
92	19.13	18.05	TMPP	Cs$_{0.05}$FA$_{0.8}$MA$_{0.15}$	1.:3	1:3	10.1002/adfm.201910710
93	21.04	18.05	TPFP	Cs$_{0.05}$FA$_{0.8}$MA$_{0.15}$	1:3	1:3	10.1002/adfm.201910710
94	21.15	20.06	2-TEAI	Cs$_{0.995}$FA$_{0.9005}$	1.003:3	0.998:3	10.1002/adma.202007431
95	21.06	18.87	QA	MA	1:3	1:3	10.1002/anie.202012095
96	18.20	17.50	BAI	Cs$_{0.0514}$FA$_{0.7905}$MA$_{0.1581}$	0.972:3	1.014:3	10.1021/acsaem.0c00553
97	20.10	17.50	BAI.F4TCNQ	Cs$_{0.0514}$FA$_{0.7905}$MA$_{0.1581}$	0.972:3	1.014:3	10.1021/acsaem.0c00553
98	21.70	20.40	TBPO	Cs$_{0.0337}$FA$_{0.6595}$MA$_{0.3067}$	1.151:3	0.924:3	10.1002/adma.201907396
99	21.20	20.40	TPPO	Cs$_{0.0337}$FA$_{0.6595}$MA$_{0.3067}$	1.151:3	0.924:3	10.1002/adma.201907396
100	21.60	18.90	ODAI$_2$	FA$_{0.8333}$MA$_{0.1667}$	0.947:3	1.026:3	10.1016/j.nanoen.2020.104892
101	20.31	19.22	HDADI	Cs$_{0.0477}$FA$_{0.8095}$MA$_{0.1428}$	0.975:3	1.012:3	10.1039/D0TA02437F
102	20.05	18.03	AVAI	MA	1.095:3	0.952:3	10.1021/acs.jpclett.0c02528
103	22.25	20.10	HBAI.FAI	Cs$_{0.05}$FA$_{0.85}$MA$_{0.1}$	0.977:3	1.012:3	10.1021/acsenergylett.0c01664
104	22.78	20.10	HBAI.FABr	Cs$_{0.05}$FA$_{0.85}$MA$_{0.1}$	0.977:3	1.012:3	10.1021/acsenergylett.0c01664
105	21.92	20.10	HBAI.FACl	Cs$_{0.05}$FA$_{0.85}$MA$_{0.1}$	0.977:3	1.012:3	10.1021/acsenergylett.0c01664
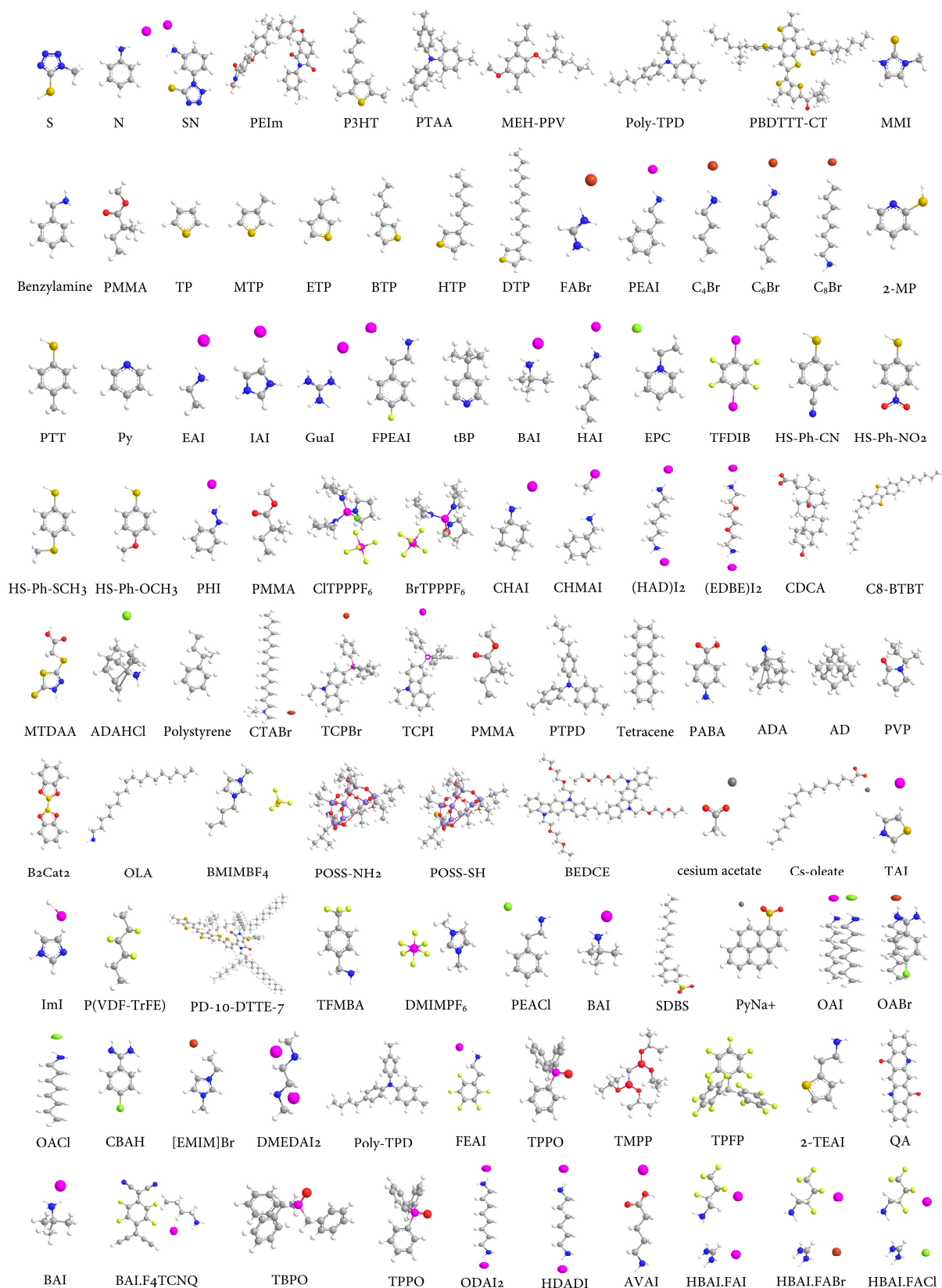
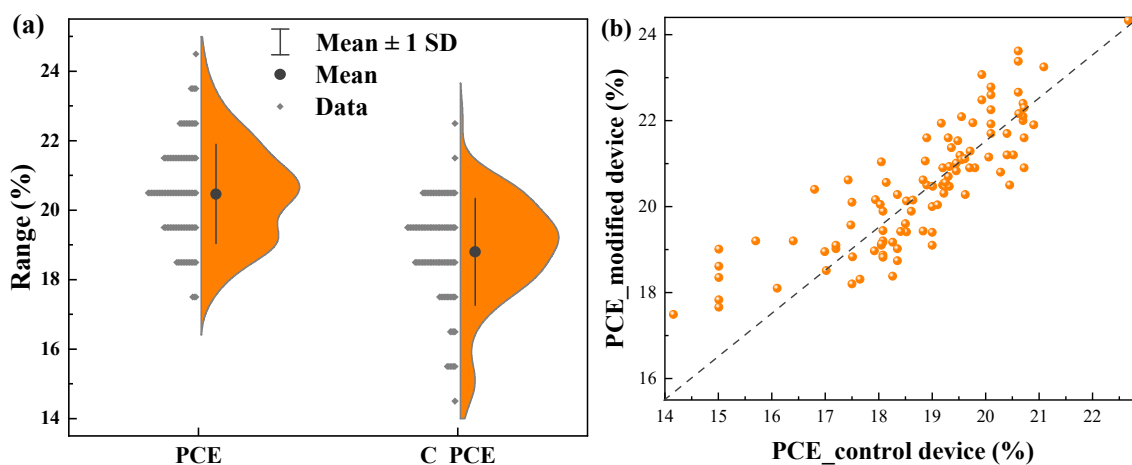**Figure S2.** Chemical structures of the interface modification materials listed in Table S1.

Figure S3. Comparison of the reported PCEs from the control devices and the modified devices by the interface modification at the perovskite/spiro-OMeTAD interface
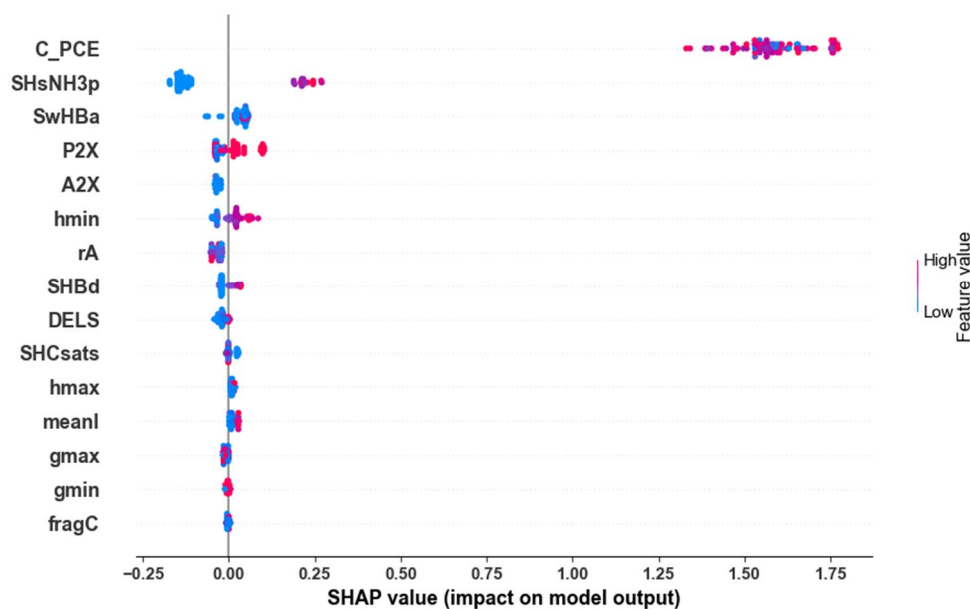

Figure S4. Distribution of the impacts (SHAP values) of the input features on the model output (Device PCE). The color represents the feature value (red high, blue low), and here only the top 15 features with the highest sum of absolute SHAP values are shown.

Table S2. RF model performance based on different combinations of the features from interface material, perovskite material, and control device performance.

Features	Train-RMSE (%)	Train-r	Test-RMSE (%)	Test-r
Interface material, control PCE	0.39	0.97	0.73	0.88
Perovskite, control PCE	0.52	0.94	0.73	0.85
Perovskite, interface material	0.57	0.94	1.10	0.49
Perovskite, interface material, control PCE	0.38	0.97	0.70	0.89

Table S3. Binding energies of the selected materials on Pb-rich surfaces of $FA_{0.6}MA_{0.4}PbI_3$.
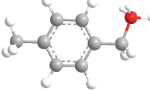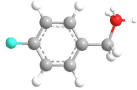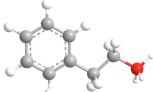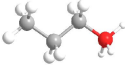
Additives	Chemical structures	Binding energy (eV)
p-MePMA$^+$		-2.19
p-F-PMA$^+$		-2.51
PEA$^+$		-2.59
PA$^+$		-3.05
EA$^+$		-3.07
MA$^+$		-3.26

Table S4. Summary of the photovoltaic parameters of the PSCs w/wo modification.

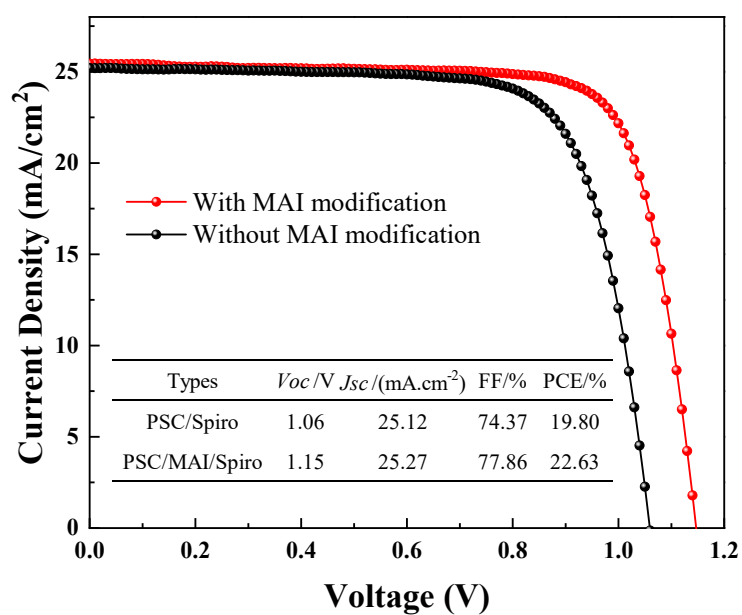| Samples	J_{SC} (mA/cm^2)	V_{OC} (V)	FF (%)	PCE (%)
Control	25.00±0.10	1.050±0.008	73.23±1.21	19.22±0.41
MAI-0.5	25.00±0.13	1.102±0.007	76.63±0.91	21.11±0.23
MAI-1.0	25.16±0.05	1.143±0.007	77.50±0.41	22.38±0.20
MAI-1.5	24.38±0.34	1.113±0.005	75.45±1.17	20.48±0.20
PEAI	25.30±0.04	1.130±0.006	76.74±0.93	21.94±0.23

Figure S5. *J-V* curves (reverse scan) and detailed parameters of the best-performed PSCs w/wo MAI modification.

Table S5. Fitted electrical parameters of PSCs based on different devices.

Samples	m_1 [a)]	m_2 [b)]	R_S (Ω cm^2)	J_0 (mA cm^{-2})
Control	2.22	2.21	1.43	2.88×10^{-6}
MAI-1.0	1.72	1.70	1.76	4.27×10^{-9}
PEAI	1.62	1.65	1.90	4.40×10^{-9}

[a)] m_1 and R_S were obtained from dV/dJ vs ($J_{SC}+J$)$^{-1}$ curve. [b)] m_2 and J_0 were obtained from ln ($J_{SC}+J$) vs ($V- R_SJ$) curve.
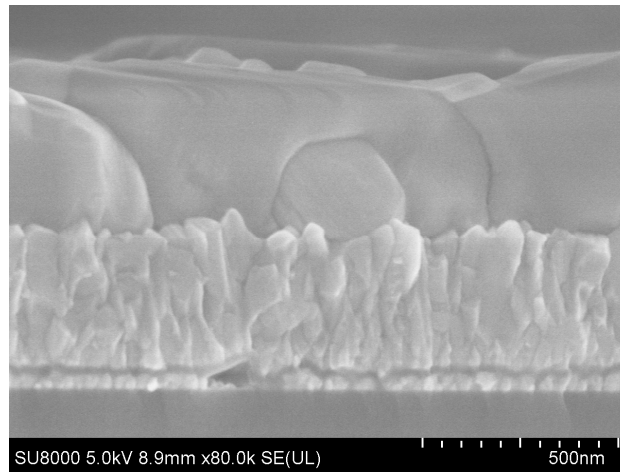

Figure S6. Cross-sectional scanning electron microscopy (SEM) image of the perovskite film on FTO for SCLC measurement. The thickness of the perovskite layer is about 500 nm.

Reference

[1] J. Shi, J. Dong, S. Lv, Y. Xu, L. Zhu, J. Xiao, X. Xu, H. Wu, D. Li, Y. Luo, Q. Meng, *Appl. Phys. Lett.,* 2014, **104**, 063901.

[2] J. You, Y. Yang, Z. Hong, T. B. Song, L. Meng, Y. Liu, C. Jiang, H. Zhou, W. H. Chang, G. Li, Y. Yang, *Appl. Phys. Lett.,* 2014, **105**, 183902.

[3] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. Probert, K. Refson, M. C. Payne, *Z. Krist.-Cryst. Mater.* 2005, **220**, 567.

[4] K. Refson, P. R. Tulip, S. J. Clark, *Phys. Rev. B*, 2006, **73**, 155114.

[5] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.,* 1996, **77**, 3865.

[6] Y. Wang, B. G. Sumpter, J. Huang, H. Zhang, P. Liu, H. Yang and H. Zhao, *J. Phys. Chem. C*, 2015, **119**, 1136.

[7] F. H. Isikgor, F. Furlan, J. Liu, E. Ugur, M. K. Eswaran, A. S. Subbiah, E. Yengel, M. D. Bastiani, G. T. Harrison, S. Zhumagali, C. T. Howells, E. Aydin, M. Wang, N. Gasparini, T. G. Allen, A. ur Rehman, E. V. Kerschaver, D. Baran, I. McCulloch, T. D. Anthopoulos, U. Schwingenschlögl, F. Laquai and S. D. Wolf, *Joule*, 2021, **5**, 1566.