

Carrier Mobility Prediction and Device Stability Evaluation of Perovskite Optoelectronic Thin Films Using Small-Sample Deep Learning Algorithms

Tai Liu^{1,*}

¹School of Physical Science and Technology, Nantong University, Nantong 226019, China

Corresponding author: Tai Liu (e-mail: liutaitai0105@163.com).

ABSTRACT Perovskite optoelectronic thin films have shown great potential in solar cells, photodetectors, light-emitting devices, and other optoelectronic applications owing to their strong light absorption, tunable bandgap, and favorable carrier transport properties. However, carrier mobility is affected by multiple factors, including material composition, processing conditions, film morphology, defect states, and testing methods, while device stability remains limited by ion migration, interfacial degradation, and environmental stress. To address the challenges of limited sample size, heterogeneous data sources, and label noise in perovskite thin-film datasets, this study develops a small-sample deep learning framework for carrier mobility prediction and device stability evaluation. A multidimensional feature system is constructed by integrating material composition, fabrication parameters, film structure, optoelectronic properties, and testing conditions. A lightweight deep learning model incorporating transfer learning, multi-task learning, regularization, and cross-validation is then developed to improve prediction reliability under limited data conditions. Stability classification, ablation experiments, and interpretability analysis are performed to identify the key factors governing carrier mobility and device degradation. The results indicate that the transfer-enhanced deep learning model improves carrier mobility prediction performance, while the multi-task learning model further enhances device stability evaluation. Feature importance analysis reveals that trap-state density, grain size, carrier lifetime, annealing temperature, and film thickness are the dominant factors affecting mobility and stability, whereas encapsulation state, humidity, and transport layer materials significantly influence long-term device stability. This study provides a data-driven approach for the rapid screening and optimization of perovskite optoelectronic thin films with high mobility and enhanced stability.

INDEX TERMS Perovskite optoelectronic thin films, Carrier mobility, Small-sample deep learning, Device stability

I. INTRODUCTION

Perovskite optoelectronic thin films have attracted extensive attention in recent years because of their strong light absorption, tunable band structure, long carrier diffusion length, and compatibility with low-temperature solution processing[1-3]. Owing to their compositional flexibility and processing versatility, perovskite materials can achieve high optoelectronic performance at relatively low fabrication costs compared with conventional inorganic semiconductors. These advantages make them promising candidates for a wide range of optoelectronic devices, including solar cells, photodetectors, light-emitting diodes, and emerging integrated photonic devices. Nevertheless, the performance of perovskite thin films is highly sensitive to material composition, crystallization quality, defect distribution, interfacial structure, and external operating conditions. In particular,

carrier mobility is a key parameter that reflects charge transport capability and directly affects carrier extraction, recombination loss, response speed, and device efficiency. Meanwhile, perovskite devices are prone to ion migration, phase segregation, moisture-induced decomposition, and interfacial degradation under humidity, thermal stress, illumination, and electric fields, which continues to limit their long-term operational stability. Therefore, accurate prediction of carrier mobility and reliable evaluation of device stability are essential for advancing the practical application of perovskite optoelectronic devices[4-6].

Current studies on carrier transport and stability in perovskite thin films mainly focus on experimental characterization, material modification, and device architecture optimization. Considerable efforts have been devoted to improving film quality and charge transport

by regulating cation/halide composition, introducing functional additives, optimizing annealing processes, controlling grain size, passivating defects, and engineering interfacial layers. Meanwhile, time-resolved photoluminescence, space-charge-limited current measurements, transient photovoltage/photocurrent techniques, Hall measurements, and impedance spectroscopy have been widely used to evaluate carrier lifetime, defect density, recombination behavior, and mobility. More recently, machine learning has been introduced into perovskite research to establish nonlinear relationships among composition, processing conditions, structural features, and device performance, showing promise in efficiency prediction, material screening, degradation analysis, and stability assessment. However, experimental datasets for perovskite thin films are usually limited in size, heterogeneous in source, inconsistent in testing conditions, and affected by label noise. As a result, conventional deep learning models may suffer from overfitting and poor generalization when directly applied to such small and heterogeneous datasets. In addition, most existing studies treat carrier mobility prediction and device stability evaluation as separate tasks. Carrier mobility is often discussed as an indicator of charge transport and extraction capability, whereas stability degradation is usually analyzed from the perspectives of environmental stress, ion migration, or interfacial instability. This separation makes it difficult to quantitatively determine whether high-mobility thin films can simultaneously maintain favorable long-term stability. Therefore, an integrated data-driven framework that links mobility prediction with stability evaluation under small-sample conditions is still needed[7-9].

In this study, a small-sample deep learning framework is proposed for carrier mobility prediction and device stability evaluation of perovskite optoelectronic thin films. First, a multidimensional feature system is constructed by integrating material composition, processing parameters, film morphology, optoelectronic properties, and testing conditions, so as to comprehensively describe the key factors governing carrier transport and stability. Second, a lightweight deep learning model suitable for small-sample scenarios is developed, together with feature selection, regularization, cross-validation, and transfer learning strategies to improve prediction reliability under limited data conditions. Furthermore, the predicted carrier mobility is combined with defect-related parameters, structural descriptors, device architecture information, and aging-performance indicators to evaluate device stability. Instead of considering mobility as the sole determinant of stability, this work treats it as an important transport-related descriptor within a broader degradation-evaluation framework. Finally, model comparison, ablation analysis, and interpretability analysis are conducted to identify the dominant factors affecting carrier mobility and stability degradation. The proposed approach is expected to provide a data-driven reference for understanding the relationship between charge transport and device aging, and to support the design and optimization of perovskite opto-

electronic thin films with both high mobility and enhanced stability.

II. MATERIALS DATA AND FEATURE CONSTRUCTION

A. DATA SOURCES AND SAMPLE DESCRIPTION

In this study, a small-sample dataset is constructed for carrier mobility prediction and device stability evaluation of perovskite optoelectronic thin films. Considering that experimental data for perovskite films are often scattered across different sources, limited in sample size, and heterogeneous in testing conditions, the data can be mainly collected from published literature, experimental records, and publicly available materials databases. To ensure comparability and model applicability, samples with relatively complete information on material composition, processing parameters, film structure, optoelectronic properties, mobility measurements, and stability indicators are preferentially selected.

The selected samples mainly focus on metal halide perovskite optoelectronic thin films, including organic-inorganic hybrid perovskites, all-inorganic perovskites, and mixed-cation/mixed-halide perovskite systems. The basic information of each sample includes perovskite composition, film fabrication method, annealing conditions, film thickness, grain size, surface roughness, optical bandgap, carrier lifetime, defect density, mobility measurement method, and device stability indicators. For different optoelectronic devices, the electron transport layer, hole transport layer, electrode material, encapsulation condition, and aging test environment are further recorded according to the device structure and operating mode[10-12]. Since this work focuses on small-sample deep learning, dataset construction does not simply pursue a larger number of samples, but emphasizes sample quality, feature completeness, and label reliability. Data with obvious outliers or unreasonable physical meanings, such as mobility values severely deviating from the typical range of similar materials, missing stability test conditions, or ambiguous definitions of efficiency retention, are removed through literature verification and consistency checking.

B. FEATURE VARIABLE CONSTRUCTION

Carrier mobility and device stability of perovskite thin films are governed by multiple interacting factors. Therefore, a multidimensional feature system is constructed to comprehensively describe the effects of material, processing, structural, and testing factors on the target properties. The selected features are mainly divided into five categories: material composition features, processing features, film morphology and structural features, optoelectronic features, and testing condition features.

Material composition features describe the crystal structure and chemical composition of perovskites. For typical ABX_3 perovskites, the A-site cation may include MA^+ , FA^+ , Cs^+ , or mixed-cation systems; the B-site metal is mainly Pb^{2+} , Sn^{2+} , or mixed Pb/Sn systems; and the X-site halide may include I^- , Br^- , Cl^- , or mixed halide ratios.

Differences in ionic radius, electronegativity, and halide composition can affect lattice distortion, bandgap, defect formation energy, and carrier transport behavior[13-15].

Processing features reflect the influence of film formation processes on crystallization quality and defect distribution. Precursor concentration, solvent system, anti-solvent type, spin-coating speed, annealing temperature, annealing time, ambient humidity, film thickness, and post-treatment strategy are selected to describe grain growth, solvent evaporation, phase transformation, and defect passivation. Film morphology and structural descriptors, including grain size, surface roughness, crystallinity, preferred orientation, trap-state density, and carrier lifetime, are extracted from SEM, AFM, XRD, PL, or TRPL characterization. Optoelectronic and testing features further include bandgap, absorption intensity, conductivity, initial efficiency, mobility measurement method, testing temperature, illumination intensity, bias voltage, aging environment, humidity, oxygen concentration, and encapsulation state.

C. DEFINITION OF TARGET LABELS

Two main research targets are defined: carrier mobility prediction and device stability evaluation. Carrier mobility is used as the first target label and can be categorized into electron mobility, hole mobility, or effective carrier mobility according to the available experimental data. Since mobility values obtained from different measurement methods may have different physical meanings, the measurement method is recorded together with the mobility value during label construction. For mobility values spanning several orders of magnitude, logarithmic transformation can be applied to reduce scale differences. Device stability is used as the second target label and can be modeled as either a regression or classification task depending on data completeness. When aging test data are sufficiently available, T80, T90, efficiency retention, performance degradation rate, or normalized stability index can be used as continuous labels. When stability data are limited or difficult to unify across experiments, device stability can be divided into high-, medium-, and low-stability levels, thereby converting stability evaluation into a classification task.

In the proposed framework, carrier mobility prediction and device stability evaluation are not treated as two completely independent tasks. Carrier mobility reflects charge transport efficiency, whereas stability reflects the ability of films and devices to maintain performance during long-term operation. Therefore, predicted mobility is combined with defect-state parameters, film morphology, aging conditions, and stability labels to establish a correlation model between mobility prediction and stability evaluation.

D. DATA PREPROCESSING AND SMALL-SAMPLE ENHANCEMENT

Because perovskite thin-film data are characterized by small sample size, high dimensionality, and strong heterogeneity, data preprocessing is a critical step for ensuring model

reliability. Continuous variables are unified in terms of units and standardized numerically. For variables with large order-of-magnitude differences, such as mobility, defect density, and resistance-related parameters, logarithmic transformation is applied to reduce the influence of extreme values during model training. Categorical variables, including A-site cation type, solvent type, additive type, device architecture, and measurement method, are encoded using one-hot encoding or embedding representations.

For missing values, different strategies are adopted according to the importance and missing ratio of each variable. Samples missing key labels are not used for the corresponding training task, while samples with a small proportion of missing input features are processed using statistical imputation, similar-sample imputation, or machine-learning-based imputation. To avoid overfitting, feature selection is performed before model construction to remove variables with high collinearity, excessive missingness, or unclear physical meaning.

For small-sample enhancement, three strategies are considered: physically reasonable data augmentation, transfer learning, and multi-task learning. Slight perturbations can be added to continuous features within the experimental error range without violating physical relationships. Transfer learning allows the model to learn general material-property relationships from related perovskite datasets before being adapted to mobility and stability prediction. Multi-task learning exploits correlations among carrier mobility, carrier lifetime, defect density, and stability labels to improve feature representation under limited data conditions.

III. SMALL-SAMPLE DEEP LEARNING METHOD

A. OVERALL RESEARCH FRAMEWORK

A small-sample deep learning framework was developed in this study to establish nonlinear mapping relationships between multidimensional features of perovskite optoelectronic thin films and two target properties, namely carrier mobility and device stability. Unlike conventional prediction approaches that rely on a single material descriptor or empirical fitting, the proposed framework integrates material composition, processing conditions, film morphology, optoelectronic properties, and testing conditions, thereby enabling a more comprehensive description of the factors governing charge transport and degradation behavior.

The overall workflow of the proposed framework is illustrated in Figure 1. First, perovskite thin-film samples are collected from literature reports, experimental records, and related datasets. Second, feature construction and pre-processing are conducted to convert heterogeneous material and device information into structured model inputs. Third, a lightweight deep learning model suitable for small-sample conditions is trained to predict carrier mobility. Fourth, the predicted mobility is combined with defect-related features, structural descriptors, and aging conditions for device stability evaluation. Finally, interpretability analysis is performed

to identify the dominant factors affecting mobility and stability.

This framework establishes a progressive relationship between mobility prediction and stability evaluation. Carrier mobility prediction focuses on the transport capability of perovskite thin films, whereas stability evaluation further considers the long-term retention of device performance under environmental and operational stress. Since device stability depends not only on initial mobility but also on ion migration, defect evolution, interfacial reactions, and external aging conditions, the two tasks are linked in a unified analytical framework rather than treated as completely independent problems.

B. FEATURE ENCODING AND INPUT REPRESENTATION

Different types of perovskite thin-film features are encoded into unified structured feature vectors. Continuous variables, such as annealing temperature, film thickness, grain size, surface roughness, bandgap, carrier lifetime, and defect density, are first checked for units and outliers. Logarithmic transformation or standardization is applied to variables that differ by several orders of magnitude. Categorical variables, such as A-site cation type, B-site metal type, solvent system, additive type, fabrication method, transport layer material, and mobility measurement method, are represented using one-hot encoding or embedding layers.

For perovskite composition features, specific chemical compositions are combined with physically meaningful descriptors, such as average ionic radius, average halide electronegativity, tolerance factor, octahedral factor, and mixed-component ratios. These descriptors reflect lattice stability, structural distortion, and defect formation tendency. Each sample is finally represented as an integrated feature vector:

$$x_i = [x_{\text{comp}}, x_{\text{proc}}, x_{\text{struc}}, x_{\text{opto}}, x_{\text{test}}] \quad (1)$$

where x_{comp} , x_{proc} , x_{struc} , x_{opto} , and x_{test} denote material composition, processing, film morphology/structure, optoelectronic, and testing-condition features, respectively.

C. CONSTRUCTION OF THE SMALL-SAMPLE DEEP LEARNING MODEL

Considering the limited size of perovskite thin-film experimental datasets, this study adopts a lightweight neural network rather than an excessively deep or heavily parameterized architecture. The model consists of an input layer, feature embedding layer, feature fusion layer, nonlinear mapping layers, and output layer. The feature embedding layer represents categorical material and processing variables, while the feature fusion layer concatenates numerical features, categorical embeddings, and physical descriptors.

In the mobility prediction task, the model output is a continuous variable representing electron mobility, hole mobility, or effective carrier mobility. Since mobility values commonly span several orders of magnitude, logarithmic

transformation is applied before modeling. The learning process can be expressed as:

$$\hat{y}_i = f_{\theta}(x_i), \quad y_i = \log(\mu_i) \quad (2)$$

The mean squared error is adopted as the basic mobility prediction loss:

$$\mathcal{L}_{\text{mobility}} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

To reduce overfitting, dropout, L2 regularization, early stopping, and K-fold cross-validation are introduced during model training. Repeated data splitting is also adopted to reduce the randomness caused by a single train-test partition.

D. TRANSFER LEARNING AND MULTI-TASK LEARNING STRATEGIES

Under small-sample conditions, training a deep learning model solely on the target dataset often leads to unstable results. Therefore, transfer learning and multi-task learning strategies are introduced to enhance feature learning and model generalization. During the pretraining stage, auxiliary tasks related to perovskite optoelectronic performance, such as bandgap prediction, absorption intensity prediction, carrier lifetime prediction, or initial device efficiency prediction, can be used to learn general material-property relationships. During fine-tuning, part of the lower feature extraction layers can be frozen to reduce overfitting.

In addition to transfer learning, multi-task learning is considered to jointly optimize mobility prediction and stability evaluation within the same model framework. The lower layers of the model are used as a shared feature extraction module, while two task-specific branches are constructed for mobility regression and stability evaluation. The total loss function is defined as:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{mobility}} + \lambda_2 \mathcal{L}_{\text{stability}} + \lambda_3 \mathcal{L}_{\text{reg}} \quad (4)$$

If stability evaluation is formulated as a regression task, $\mathcal{L}_{\text{stability}}$ can be defined as the mean squared error. If it is formulated as a classification task, cross-entropy loss can be used. Through joint training, the model exploits complementary information among related tasks and learns more robust feature representations under limited data conditions.

E. STABILITY EVALUATION AND INTERPRETABILITY ANALYSIS

Device stability evaluation is formulated in two ways: continuous lifetime prediction and stability-level classification. When aging test data are sufficiently complete, T80, T90, efficiency retention, or degradation rate is used as a continuous label. When aging conditions differ significantly among samples or stability data are incomplete, samples are divided into high-, medium-, and low-stability levels according to efficiency retention and test duration. To avoid treating the deep learning model as a black box, interpretability analysis is introduced to reveal the influence of different features on

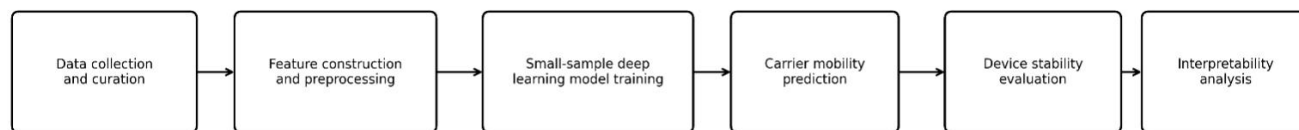


FIGURE 1. Overall workflow of the proposed small-sample deep learning framework.

carrier mobility and device stability. SHAP values, permutation importance, and partial dependence analysis can be used to evaluate feature contributions. Through interpretability analysis, dominant factors such as trap-state density, grain size, carrier lifetime, annealing temperature, encapsulation state, humidity, and transport layer material can be linked with material physics mechanisms.

F. ALGORITHMIC PROCEDURE

To clearly describe the implementation process of the proposed method, the algorithmic procedure is summarized in Figure 2. The workflow starts from the input of perovskite thin-film samples, followed by multidimensional feature extraction, including composition, processing, structure, optoelectronic, and testing-condition features. After numerical normalization and categorical encoding, the dataset is divided using repeated K-fold cross-validation to improve the robustness of model evaluation under small-sample conditions. A lightweight deep learning model is then trained with regularization strategies such as dropout, L2 regularization, and early stopping. After model training, carrier mobility is predicted on the logarithmic scale. The predicted mobility is subsequently combined with degradation-related indicators, such as defect density, carrier lifetime, grain size, encapsulation state, humidity, and aging conditions, for device stability evaluation. Finally, SHAP-based interpretability analysis is conducted to quantify the contribution of different features to mobility prediction and stability classification. Compared with a purely sequential prediction pipeline, this procedure enables the model to exploit the correlation between transport-related features and degradation-related features. As a result, the framework is not only capable of estimating carrier mobility under limited data conditions, but also suitable for evaluating device stability in a more integrated and physically meaningful manner.

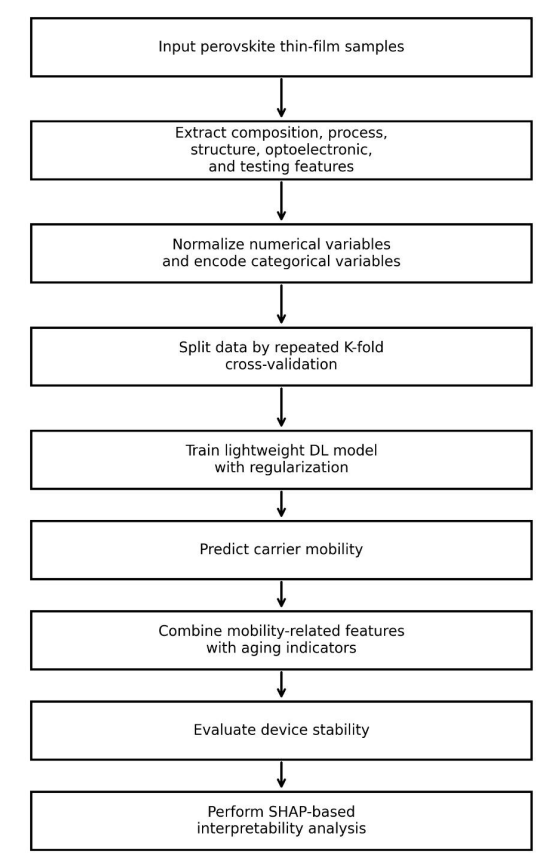


FIGURE 2. Algorithmic procedure for mobility prediction and stability evaluation.

TABLE 1. Performance comparison of different models for carrier mobility prediction

Model	R^2	RMSE_log	MAE_log
Linear regression	0.58	0.61	0.47
SVR	0.67	0.52	0.40
Random forest	0.74	0.45	0.34
XGBoost	0.79	0.39	0.30
Small-sample DL	0.82	0.36	0.27
Transfer-enhanced DL	0.87	0.31	0.23

learning models achieved improved performance, suggesting that complex interactions among material composition, processing parameters, and film structure are important for mobility prediction. The proposed small-sample deep learning model achieved higher R^2 and lower errors, and the transfer-enhanced model further improved R^2 to 0.87, indicating that pretraining helps the model learn more general material-property relationships. Figure 3 shows the relationship be-

IV. RESULTS AND DISCUSSION

A. CARRIER MOBILITY PREDICTION PERFORMANCE

To validate the effectiveness of the proposed model, it was compared with linear regression, support vector regression, random forest, and XGBoost. Since carrier mobility values usually span several orders of magnitude, model training was conducted on the logarithmic scale. Table I summarizes the prediction performance of different models.

As shown in Table I, the linear regression model exhibited relatively limited predictive performance, indicating that the relationship between perovskite thin-film features and carrier mobility is not simply linear. Nonlinear machine

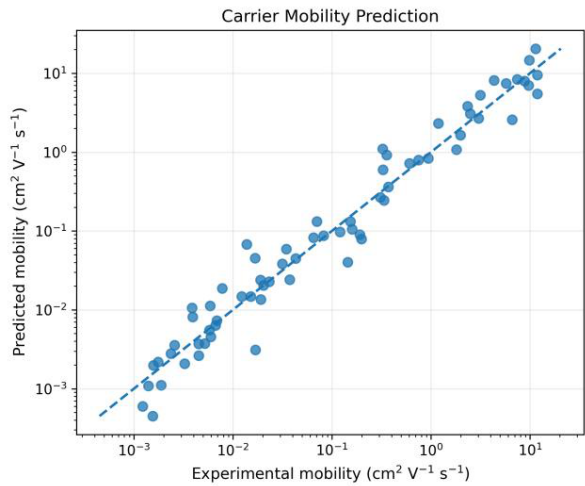


FIGURE 3. Predicted and experimental carrier mobility values of perovskite optoelectronic thin films.

TABLE 2. Performance comparison of different models for device stability classification

Model	Accuracy	Precision	Recall	F1-score
Random forest	0.76	0.75	0.74	0.74
XGBoost	0.80	0.79	0.78	0.78
Small-sample DL	0.83	0.82	0.81	0.81
Multi-task DL	0.88	0.87	0.86	0.86

tween predicted and experimental carrier mobility values obtained by the transfer-enhanced deep learning model. Most data points are distributed near the diagonal line, indicating good agreement between the predicted and experimental values. The relatively large deviations observed for some samples may be related to differences in measurement methods, heterogeneous defect distributions, or label noise in literature-derived data.

B. DEVICE STABILITY EVALUATION RESULTS

Based on carrier mobility prediction, device stability was further evaluated by integrating defect density, carrier lifetime, grain size, transport layer material, encapsulation state, and aging test conditions. Considering that stability testing conditions differ among reported studies, stability evaluation was formulated as a classification task. Table II presents the classification performance of different models.

The results show that the multi-task deep learning model achieved the best stability classification performance, with an accuracy of 0.88 and an F1-score of 0.86. Compared with the single-task model, the multi-task model simultaneously exploited shared information between mobility prediction and stability classification, thereby improving its ability to identify features related to device degradation. Figure 4 indicates that high-stability and low-stability samples were identified with relatively high accuracy, whereas medium-stability samples were more likely to be confused with adjacent classes. This is because medium-stability devices are usually located in a transitional degradation region,

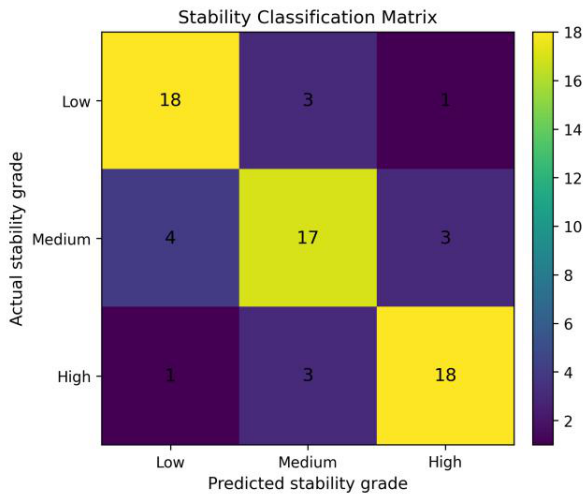


FIGURE 4. Confusion matrix of the multi-task deep learning model for device stability classification.

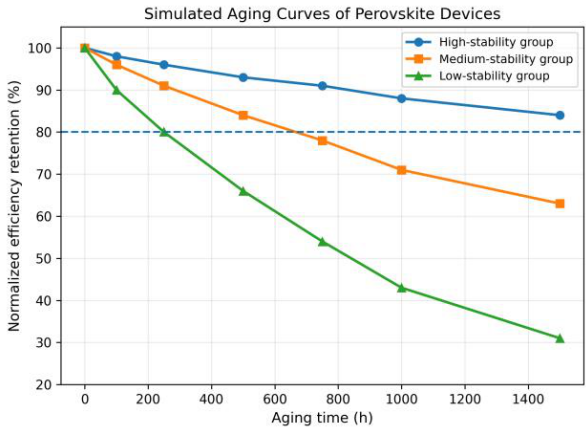


FIGURE 5. Normalized efficiency retention curves of perovskite devices with different stability levels.

where aging behavior is jointly affected by encapsulation quality, humidity, interfacial stability, and defect passivation.

C. STABILITY DEGRADATION BEHAVIOR

To further analyze the degradation behavior of devices with different stability levels, normalized efficiency retention curves were constructed, as shown in Figure 5. The high-stability group retained approximately 84% of its initial performance after 1500 h of aging, indicating favorable film structure and interfacial stability. The medium-stability group gradually decreased below 80% retention after approximately 750–1000 h, showing progressive degradation. In contrast, the low-stability group exhibited obvious early-stage performance decay.

The degradation curves indicate that stability differences are reflected not only in final efficiency retention but also in degradation rate and degradation onset time. For low-stability devices, rapid early-stage degradation may be associated with unpassivated defects, active ion migration,

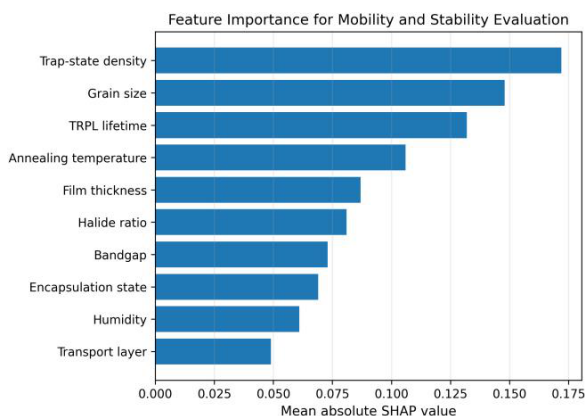


FIGURE 6. Feature importance ranking based on mean absolute SHAP values.

moisture and oxygen penetration along grain boundaries, and interfacial reactions with transport layers. Therefore, stability evaluation should integrate aging time, retention variation, and environmental conditions rather than rely solely on a single endpoint indicator.

D. FEATURE IMPORTANCE AND MECHANISTIC INTERPRETATION

SHAP analysis was performed to quantify the contribution of different features to mobility prediction and stability evaluation. Figure 6 shows the mean absolute SHAP values of the major features. Trap-state density, grain size, TRPL carrier lifetime, annealing temperature, and film thickness were identified as the most influential factors.

Trap-state density showed the highest importance, indicating that defect-related recombination and trap-assisted transport strongly influence both carrier mobility and device stability. Higher trap-state density increases the probability of carrier trapping, reduces effective mobility, promotes non-radiative recombination, and may accelerate local ion migration. Grain size ranked as the second most important feature, suggesting that grain-boundary density plays a critical role in carrier transport and stability. Larger grains generally indicate fewer grain-boundary defects and more continuous transport pathways. TRPL carrier lifetime also contributed significantly, demonstrating that photogenerated carrier recombination dynamics are closely related to mobility and stability. Annealing temperature and film thickness reflect the influence of fabrication conditions on crystallization quality and charge transport pathways. Encapsulation state, humidity, and transport layer material were not the most dominant factors for mobility prediction, but they made considerable contributions to stability evaluation, indicating that long-term stability depends strongly on external protection, interface engineering, and aging environment.

E. ABLATION AND ROBUSTNESS ANALYSIS

Ablation experiments were conducted to verify the contribution of different feature groups. While keeping the model

TABLE 3. Ablation analysis of different feature groups

Feature setting	Mobility R^2	Stability F1-score
All features	0.87	0.86
Without composition features	0.80	0.81
Without processing features	0.78	0.79
Without morphology/structure features	0.71	0.73
Without optoelectronic features	0.75	0.76
Without testing-condition features	0.82	0.80

structure and training strategy unchanged, material composition features, processing features, morphology/structural features, optoelectronic features, and testing-condition features were removed separately. Table III presents the resulting changes in mobility prediction and stability evaluation performance.

Removing morphology and structural features caused the most significant performance degradation. The mobility prediction R^2 decreased from 0.87 to 0.71, while the stability F1-score decreased from 0.86 to 0.73. This indicates that grain size, roughness, crystallinity, and trap-state density serve as key bridges connecting material composition, processing conditions, and final device performance. Removing testing-condition features caused a relatively smaller decline in mobility prediction performance but a more evident decrease in stability evaluation performance, confirming the importance of aging conditions for stability analysis.

Overall, the results demonstrate that a multidimensional feature system is necessary for reliable prediction and evaluation. Relying solely on material composition or initial optoelectronic performance is insufficient to explain variations in mobility and stability. A more reliable model can only be established by jointly considering composition, processing, structure, optoelectronic properties, and testing conditions.

F. RELATIONSHIP BETWEEN CARRIER MOBILITY AND DEVICE STABILITY

Although carrier mobility and device stability are both critical indicators for evaluating perovskite optoelectronic thin films, their relationship is not strictly linear or monotonic. In general, higher carrier mobility indicates more efficient charge transport, faster carrier extraction, and lower probability of bulk recombination, which are beneficial for improving the initial performance of optoelectronic devices. For solar cells, enhanced mobility can reduce charge accumulation and suppress recombination loss at the interface; for photodetectors, it can improve response speed and carrier collection efficiency; and for light-emitting devices, balanced carrier transport can contribute to more efficient radiative recombination. Therefore, carrier mobility can be regarded as an important transport-related descriptor in device-performance prediction.

However, high carrier mobility does not necessarily guarantee long-term device stability. Device stability is governed by a broader set of factors, including defect evolution, ion migration, phase segregation, interfacial reaction, moisture sensitivity, thermal stress, and light-induced degradation. A

perovskite film may exhibit relatively high mobility in the initial state, but if its grain boundaries are not sufficiently passivated or its interface is chemically unstable, the device may still experience rapid performance decay during aging. For example, unpassivated grain boundaries can provide fast pathways for ion migration and moisture penetration, while unstable transport layers may accelerate interfacial decomposition under illumination or thermal stress. In such cases, the initial carrier transport capability cannot fully represent the long-term reliability of the device.

The results of this study also support this mechanism. Features such as trap-state density, grain size, carrier lifetime, annealing temperature, and film thickness show strong contributions to both mobility prediction and stability evaluation, indicating that film crystallization quality and defect characteristics serve as common factors linking charge transport and degradation behavior. Nevertheless, encapsulation state, humidity, and transport layer material contribute more significantly to stability evaluation than to mobility prediction. This suggests that stability is not only determined by intrinsic film transport properties, but also strongly affected by external environmental conditions and interfacial protection strategies.

Therefore, in the proposed framework, carrier mobility is not treated as the only determinant of device stability, but rather as an important intermediate descriptor connecting film microstructure and device aging behavior. The combination of mobility-related features with degradation-related variables enables a more comprehensive evaluation of perovskite devices. From a materials-design perspective, achieving high-performance and stable perovskite optoelectronic devices requires the simultaneous optimization of charge transport, defect passivation, crystallization quality, interfacial compatibility, and environmental resistance. This also explains why a joint modeling strategy that integrates mobility prediction and stability evaluation is more informative than a single-task model focused only on initial transport properties.

V. CONCLUSION

This study developed a small-sample deep learning framework for carrier mobility prediction and device stability evaluation of perovskite optoelectronic thin films. By integrating multidimensional features, including material composition, processing parameters, film morphology, optoelectronic properties, and testing conditions, the proposed framework established a relationship from film structural characteristics to carrier transport behavior and further to device stability. The results show that the lightweight small-sample deep learning model can effectively capture the nonlinear relationship between perovskite thin-film features and carrier mobility. Moreover, the incorporation of transfer learning and multi-task learning further improves prediction accuracy and stability evaluation performance. The interpretability analysis indicates that trap-state density, grain size, carrier lifetime, annealing temperature, and film thick-

ness are the dominant factors affecting carrier mobility and device stability. These findings suggest that the optimization of perovskite devices should not rely solely on mobility enhancement, but should also consider crystallization quality, defect passivation, interfacial stability, and environmental protection.

Future work may be further extended by incorporating more experimental data and standardized aging-test results, integrating multimodal characterization data such as SEM, AFM, XRD, PL, and in situ aging images, and introducing physical constraints, active learning, and Bayesian uncertainty estimation. With the development of high-quality materials databases and intelligent optimization algorithms, small-sample deep learning is expected to become an effective tool for rapid screening and optimization of perovskite optoelectronic thin films with high mobility and long-term stability.

DECLARATION OF CONFLICTING INTERESTS The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

DATA SHARING AGREEMENT The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

FUNDING The author(s) received no financial support for the research, authorship, and/or publication of this article.

REFERENCES

- [1] Q. L. Tao, P. C. Xu, M. J. Li, et al., "Machine learning for perovskite materials design and discovery," *npj Comput. Mater.*, vol. 7, p. 23, 2021.
- [2] V. De La Asuncion-Nadal et al., "Machine learning for perovskite solar cells," *Energy Adv.*, 2025.
- [3] J. Wang, X. Liu, Y. Yang, et al., "Feature selection in machine learning for perovskite materials," *Materials*, vol. 16, no. 8, p. 3134, 2023.
- [4] Z. Zhang, H. H. Wang, J. Jacobsson, et al., "Big data driven perovskite solar cell stability analysis," *Nat. Commun.*, vol. 13, p. 7639, 2022.
- [5] D. Zhang, M. H. Kumar, N. G. Park, et al., "Degradation pathways in perovskite solar cells and how to meet international standards," *Commun. Mater.*, vol. 3, p. 58, 2022.
- [6] S. Baumann, D. Barrit, J. Schall, et al., "Stability and reliability of perovskite containing solar cells and modules: Degradation mechanisms and mitigation strategies," *Energy Environ. Sci.*, vol. 17, no. 20, pp. 7358–7400, 2024.
- [7] I. Kouroudis, V. Paraskeva, A. Livera, et al., "Artificial intelligence-based, wavelet-aided prediction of long-term outdoor performance of perovskite solar cells," *ACS Energy Lett.*, vol. 9, pp. 1888–1898, 2024.
- [8] S. S. Zhao, S. J. Zhou, Z. L. Guo, et al., "Machine learning-assisted analysis of perovskite solar cell long-term stability under multiple environmental factors," *ACS Sustainable Chem. Eng.*, vol. 13, no. 19, pp. 7300–7311, 2025.
- [9] W. A. Dunlap-Shohl, Y. Meng, P. P. Sunkari, et al., "Physiochemical machine learning models predict operational lifetimes of $\text{CH}_3\text{NH}_3\text{PbI}_3$ perovskite solar cells," *J. Mater. Chem. A*, vol. 12, pp. 9730–9741, 2024.

- [10] B. N. N. Alsulami et al., "Application of large datasets to assess trends in perovskite solar cell stability using machine learning," *J. Mater. Chem. A*, vol. 12, pp. 10172–10184, 2024.
- [11] J. Lim, M. Kober-Czerny, Y. H. Lin, et al., "Long-range charge carrier mobility in metal halide perovskite thin-films and single crystals via transient photo-conductivity," *Nat. Commun.*, vol. 13, p. 4201, 2022.
- [12] Z. Zhang and B. Saparov, "Charge carrier mobility of halide perovskite single crystals for ionizing radiation detection," *Appl. Phys. Lett.*, vol. 119, no. 3, p. 030502, 2021.
- [13] V. M. Le Corre, E. A. Duijnste, O. El Tambouli, et al., "Revealing charge carrier mobility and defect densities in metal halide perovskites via space-charge-limited current measurements," *ACS Energy Lett.*, vol. 6, no. 3, pp. 1087–1094, 2021.
- [14] V. Gupta, K. Choudhary, F. Tavazza, et al., "Cross-property deep transfer learning framework for enhanced predictive analytics on small materials data," *Nat. Commun.*, vol. 12, p. 6595, 2021.
- [15] V. Gupta, K. Choudhary, F. Tavazza, et al., "Structure-aware graph neural network based deep transfer learning framework for enhanced predictive analytics on diverse materials properties," *npj Comput. Mater.*, vol. 10, p. 5, 2024.