Research Article | Open Access

Journal of Materials Informatics

Chen *et al. J. Mater. Inf.* **2026**, *6*, 28 **DOI:**10.20517/jmi.2025.76

Machine learning-driven morphology identification and classification of high-throughput functional oxide films

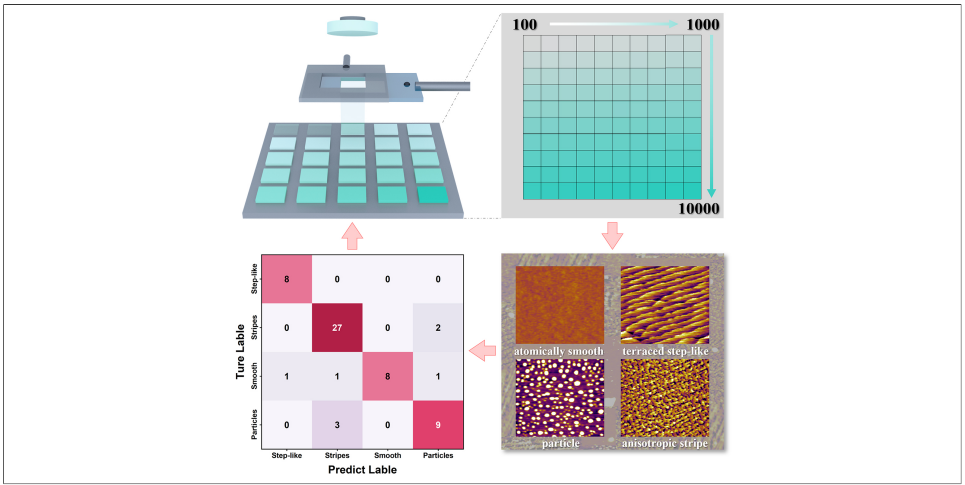
Qianxin Chen[1,2,3,#], Xiangdong Wang[1,#], Liyufen Dai[1,2], Hongjia Song[1], Jinbin Wang[1], Xiangli Zhong[1,*], Juan Zou[4,*], Gaokuo Zhong[2,3,*]

Keywords:
Machine learning, atomic force microscopy, functional oxide films, $SrRuO_3$, automated identification and classification

Citation: Chen, Q.; Wang, X.; Dai, L.; Song, H.; Wang, J.; Zhong, X.; Zou, J.; Zhong, G. Machine learning-driven morphology identification and classification of high-throughput functional oxide films. *J. Mater. Inf.* **2026**, *6*, 28. https://dx.doi.org/10.20517/jmi.2025.76

Received: 5 Sep 2025 **First Decision:** 23 Oct 2025 **Revised:** 3 Nov 2025 **Accepted:** 18 Nov 2025 **Published:** 29 Apr 2026

Academic Editors: Sergei Manzhos, Hao Li **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang



Abstract

Functional oxide films offer precise control over diverse properties through tunable physical characteristics and interface effects, with their functionality primarily determined by morphology. However, conventional methods are incapable of obtaining large-scale morphological data and face significant challenges in data identification and classification, which fundamentally limit the rapid assessment of thin film properties and functional screening. Herein, we establish a comprehensive morphological database of oxide films utilizing high-throughput experimental methods and develop a machine learning framework for automated identification and classification of atomic force microscopy data. Using gradient-thickness $SrRuO_3$ films as a representative example, this framework achieves enhanced performance through hyperparameter optimization and strategic adjustments, ultimately reaching a classification accuracy of 86.67% in independent tests, demonstrating its effectiveness in morphology analysis of functional oxide films.

[1]National-Provincial Laboratory of Special Function Thin Film Materials, School of Materials Science and Engineering, Xiangtan University, Xiangtan 411105, Hunan, China.
[2]Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China.
[3]Changsha Semiconductor Technology and Application Innovation Research Institute, College of Semiconductors (College of Integrated Circuits), Hunan University, Changsha 410082, Hunan, China.
[4]School of Artificial Intelligence, Changsha University of Science and Technology, Changsha 410114, Hunan, China.
[#]These authors contributed equally to this work.

[*]**Correspondence to:** Prof. Gaokuo Zhong, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, Guangdong, China. E-mail: gkzhong@hnu.edu.cn; Dr. Juan Zou, School of Artificial Intelligence, Changsha University of Science and Technology, Changsha 410114, Hunan, China. E-mail: zjuan@csust.edu.cn; Prof. Xiangli Zhong, National-Provincial Laboratory of Special

Function Thin Film Materials, School of Materials Science and Engineering, Xiangtan University, Xiangtan 411105, Hunan, China. E-mail: xlzhong@xtu.edu.cn

Furthermore, this approach shows significant potential for automated microstructure analysis of complex oxides and is expected to accelerate research on structure-property correlations in functional oxide films.

INTRODUCTION

Functional oxide films have emerged as a critical class of materials in advanced technology sectors due to their highly tunable physical properties, compatibility with conventional semiconductor processes, and distinctive interface-mediated phenomena[1-5]. These characteristics collectively facilitate enhanced and frequently novel nanoscale functionalities, such as ferroelectric switching[6-10], piezoelectric responses[11,12], multiferroic behavior[13,14], and cross-coupling effects[15], thereby establishing functional oxide films as indispensable components across modern microelectronics, optoelectronics, energy conversion systems, and information storage technologies[16]. The quality and performance orientation of films can be comprehensively evaluated through a suite of optical and electrical characterization techniques such as atomic force microscopy (AFM)[17], X-ray diffraction (XRD)[18], and reciprocal space mapping (RSM)[19]. Among these, film morphology represents one of the most fundamental and intuitive types of data, critically influencing crystallographic quality and functional stability[20]. Thus, establishing dynamic relationships among processing conditions, morphology, and functional properties is essential for rapidly guiding experiments and screening desired performance. However, current research still relies heavily on iterative trial-and-error methods based on individually fabricated samples[21]. This approach is not only time-consuming and labor-intensive but also prone to experimental variability, particularly for complex functional oxides highly sensitive to processing parameters[21]. Moreover, optimization derived from limited sample sizes and compositional ranges may lead to missed opportunities for discovering novel physical phenomena or achieving optimal functional properties at specific components[22].

In recent years, the multidisciplinary Materials Genome Initiative (MGI) has attracted broad attention for its ability to accelerate the materials development cycle and substantially reduce associated time and costs through integrated computational materials design and performance prediction[23-26]. Central to this effort is the continuous improvement of high-throughput experimentation (HTE) and high-throughput characterization (HTC) techniques, which constitute essential components of the MGI framework[27,28]. HTE combines thin-film deposition with fast-moving masks, enabling highly efficient parallel or sequential experimentation on individual processing parameters[22,29]. This methodology has been successfully employed to design and fabricate high-throughput (HT) films with precisely controlled thickness[29,30], composition[22], and structural gradients[31]. To accelerate the structural characterization of HT samples, HTC techniques have emerged[28]. These methods leverage database screening and rapid feedback mechanisms to significantly shorten the structure-property analysis and research cycle[28]. Novel HTC approaches have been developed based on AFM[32,33], combinatorial XRD[34,35], and second harmonic generation (SHG)[28,36]. The integration of HTE and HTC holds significant potential for rapidly constructing extensive databases of functional oxide films and accelerating performance screening, feedback, and optimization processes. However, these methods primarily focus on broad parameter exploration rather than targeted structural and property modulation, which fundamentally limits their capacity for deterministic materials design[37]. To address these challenges, we propose integrating a machine learning (ML)-based artificial intelligence framework with MGI to further accelerate the establishment of reliable structure-property relationships. Based on composition-gradient samples prepared through HTE, microstructural data and property data are systematically collected via HTC to construct standardized databases. The ML framework is then employed to automatically extract microstructural features and transform them into quantifiable feature vectors, while feature engineering techniques are used to identify critical material descriptors. Furthermore, supervised

learning algorithms are applied to establish nonlinear mapping models that connect material descriptors and structural features to target properties. This approach enables more effective prediction of performance for novel compositions or structures, facilitates the establishment of more robust structure-property relationships, and achieves accurate prediction and optimized design of material properties[38-42], thereby significantly accelerating the development cycle of novel functional oxide thin films. It is noteworthy that different functional oxide materials possess distinct material characteristics and functional properties. Therefore, ML approaches should be appropriately adjusted and refined according to these specific material features to significantly enhance the learning capability of the framework. Such tailored optimization is expected to further improve the reliability of structure-property relationships across different oxide material systems.

In this study, $SrRuO_3$ (SRO) thin films were selected as a representative model system for functional oxides due to their high electrical conductivity, structural stability, and excellent process compatibility. Based on thickness-gradient SRO films fabricated via HTE, we developed an automated AFM image recognition and classification framework integrating preprocessing algorithms and ML. Through hyperparameter optimization and strategic tuning, the model achieved a classification accuracy of 86.67% in independent tests, successfully establishing correlations between deposition parameters and film morphology, thereby enabling predictive modeling for functional oxide thin films.

MATERIALS AND METHODS

Fabrication of high-throughput SRO thin films

Gradient-thickness SRO films were deposited using high-throughput pulsed laser deposition (RP-HT-102, Shenzhen Arrayed Materials Co., Ltd., China) with a 248 nm KrF excimer laser. A 10 × 10 gradient thickness array comprising 100 regions was fabricated, wherein the number of laser pulses varied from 100 to 10,000. The substrate temperature and oxygen pressure were maintained at 690 °C and 100 mTorr, respectively.

Characterization of structural properties

Topographical analyses were conducted using AFM (Cypher ES, Asylum Research) with a conductive probe (e.g., ZDYF2022-EC450, Institute of Semiconductors, Chinese Academy of Sciences).

RESULTS AND DISCUSSION

Our integrated experimental framework is illustrated in Figure 1A. Based on high-throughput pulsed laser deposition system (HT-PLD), we utilized a movable mask within the system to fabricate continuously varying gradient-thickness SRO films with a 10 × 10 array on an $SrTiO_3$ (STO) substrate by systematically varying the number of deposition pulses (from 100 to 10,000 shots). Full-area optical microscopy images revealed no significant chromatic aberration across the visible spectrum for SRO films with different thicknesses, demonstrating the relative stability of their optical properties against thickness variations. To further investigate the microstructural evolution, we characterized the surface morphology of gradient-thickness SRO films using AFM, as shown in Figure 1B. Distinct morphological differences were observed, which can be broadly categorized into four types: atomically smooth morphologies, particle morphologies, anisotropic stripe morphologies, and terraced step-like morphologies. The detailed correspondence between the number of deposition pulses and surface morphology is systematically documented in Supplementary Table 1. This well-defined relationship established a reliable foundation for the ML framework to develop accurate mappings from deposition parameters to surface morphology. We further developed an automated image processing workflow based on Python to minimize non-structural interference in the morphological data. This method primarily incorporates an adaptive edge detection algorithm, which effectively eliminates scale bars and color legends to a certain extent. The preprocessed AFM images [Figure 1C] provide a reliable foundation for ML-driven morphological classification. To

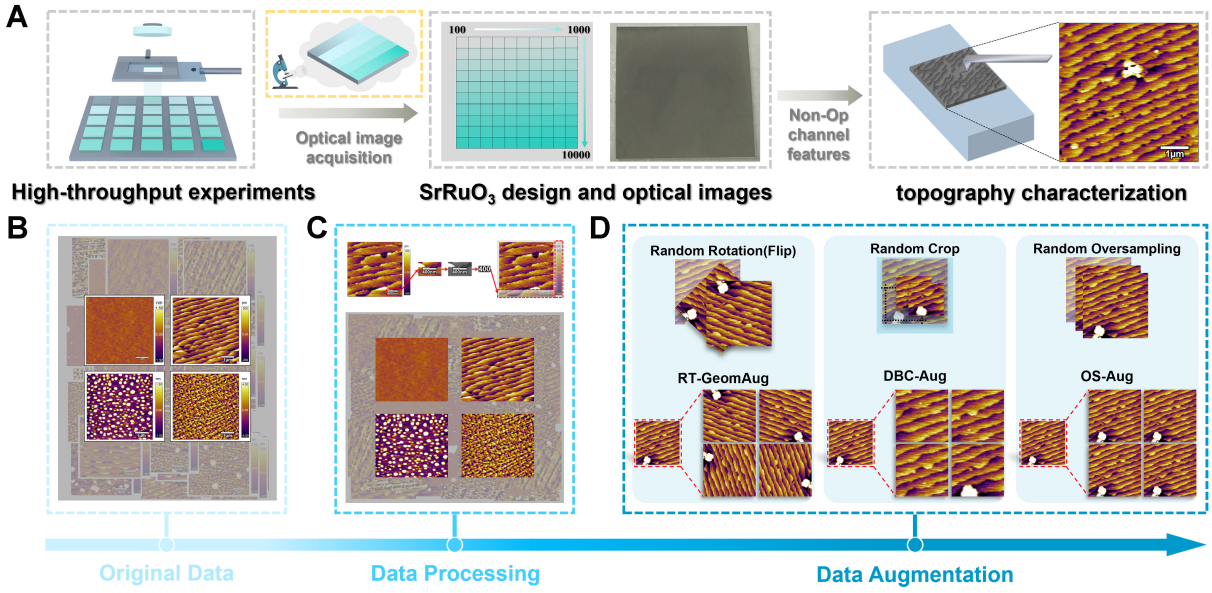


Figure 1. Preparation, characterization, and data processing of gradient-thickness SRO films. (A) Schematic diagram of the preparation and characterization process for gradient-thickness SRO films; (B) Original AFM surface morphology of gradient-thickness SRO films; (C and D) Morphologies after (C) automatic preprocessing and (D) processing with different data augmentation strategies (RT-GeomAug, DBC-Aug, and OS-Aug). SRO: $SrRuO_3$; AFM: atomic force microscopy; RT-GeomAug: real-time geometric augmentation; DBC-Aug: dynamic balance cropping augmentation; OS-Aug: over-sampling augmentation.

address the challenges of limited dataset size and class imbalance in morphological images of SRO films, we implemented three complementary data augmentation strategies, including real-time geometric augmentation (RT-GeomAug, random rotations within ± 180° combined with horizontal/vertical flipping applied to raw morphology images), dynamic balance cropping augmentation (DBC-Aug; modulates category-specific cropping frequency to equalize sample distribution) and over-sampling augmentation (OS-Aug; repeated sampling cycles until all classes reached equivalent maximum counts). These synergistic approaches were systematically optimized to identify the most effective data processing pipeline for our specific dataset, with the enhanced morphological features clearly demonstrated in Figure 1D.

Based on the preprocessed AFM images, we systematically trained and evaluated different ML models following the workflow shown in Figure 2A. The input AFM images first underwent standardized resizing and normalization procedures to ensure dimensional consistency and data accuracy, wherein the dataset was partitioned into training, validation, and test sets at a ratio of 7:1.5:1.5 [Supplementary Table 2]. Notably, the ratio was determined through multiple rounds of computational adjustment and validation; excessively high or low proportions of training samples may compromise evaluation reliability or lead to underfitting. Furthermore, under the three different data augmentation strategies, the sizes of the validation and test sets were kept consistent to ensure the reliability of evaluation results. The number of samples in the training, validation, and test sets under each augmentation strategy is detailed in Supplementary Tables 3-5. The sample size used in this study was determined based on the thickness-gradient SRO oxide system. For functionally or structurally more complex oxide systems, the framework requires larger and more diverse datasets incorporating different characterization techniques, along with corresponding adjustments and improvements tailored to specific material systems. The preprocessed images were then classified through convolutional neural network (CNN), support vector machine (SVM), and k-nearest neighbor (KNN) models. Notably, while the CNN model could directly process and classify the preprocessed images, both SVM and KNN models required feature extraction through a ResNet-18 architecture to generate 128-dimensional feature vectors as input representations. To ensure a valid comparative evaluation, identical

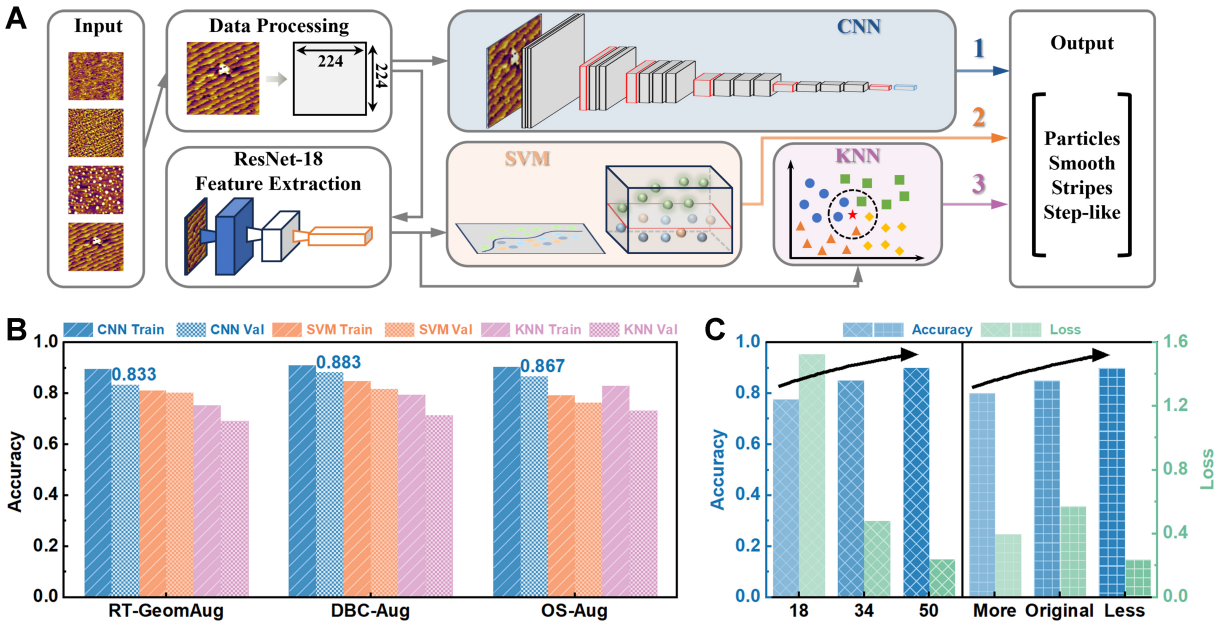


Figure 2. ML model architecture design and optimization strategies. (A) Schematic workflow of the computational framework for training and evaluating various ML models; (B) Performance comparison of different classification models under three data augmentation strategies; (C) Evaluation of classification accuracy and loss function trends across ResNet architectures with varying depths and multi-level feature freezing strategies. ML: Machine learning; CNN: convolutional neural network; SVM: support vector machine; KNN: k-nearest neighbor; RT-GeomAug: real-time geometric augmentation; DBC-Aug: dynamic balance cropping augmentation; OS-Aug: over-sampling augmentation.

training-validation-test splits were maintained across all models, and consistent evaluation metrics were applied throughout. We compared the accuracy of the ML models under different data augmentation strategies, as shown in Figure 2B. The plot clearly demonstrates that the CNN model achieved higher accuracy than both SVM and KNN models on both the training and validation sets across all augmentation strategies. Specifically, the validation accuracy of the CNN model reached 83.30%, 88.30%, and 86.70% under the RT-GeomAug, DBC-Aug, and OS-Aug strategies, respectively. Consequently, the CNN model with the DBC-Aug strategy was selected for subsequent validation. To further improve performance, ResNet architectures were incorporated into the CNN model to optimize both architectural configurations and feature freezing strategies [Figure 2C]. First, CNN models with ResNet architectures of varying depths were evaluated, and the ResNet-50 configuration delivered the best performance with 90.00% accuracy. Building on this ResNet-50 architecture, different feature freezing strategies were then tested. Comparative analysis revealed that the "Less" freezing strategy outperformed both the "More" and "Original" approaches. Through the systematic implementation of these two optimization strategies, a final CNN model was developed that combines the DBC-Aug data augmentation strategy with a ResNet-50 architecture employing the "Less" feature freezing approach. This optimized configuration achieved peak classification accuracy, demonstrating superior performance in comparative analyses. Based on these results, this model configuration was adopted as the primary ML framework for all subsequent morphology image recognition and classification tasks, serving as the benchmark solution for the AFM image analysis pipeline.

The workflow for SRO morphological classification is illustrated in Figure 3A. Our approach employs a CNN classification model based on the ResNet-50 architecture, using three-channel grayscale images as input throughout the process. The network retains the first four convolutional blocks of ResNet-50, while optimizing the freezing strategy and feature transfer mechanism within the framework, specifically by unfreezing the model parameters from layers 2 to 4. This architecture stems from an analysis of hierarchical

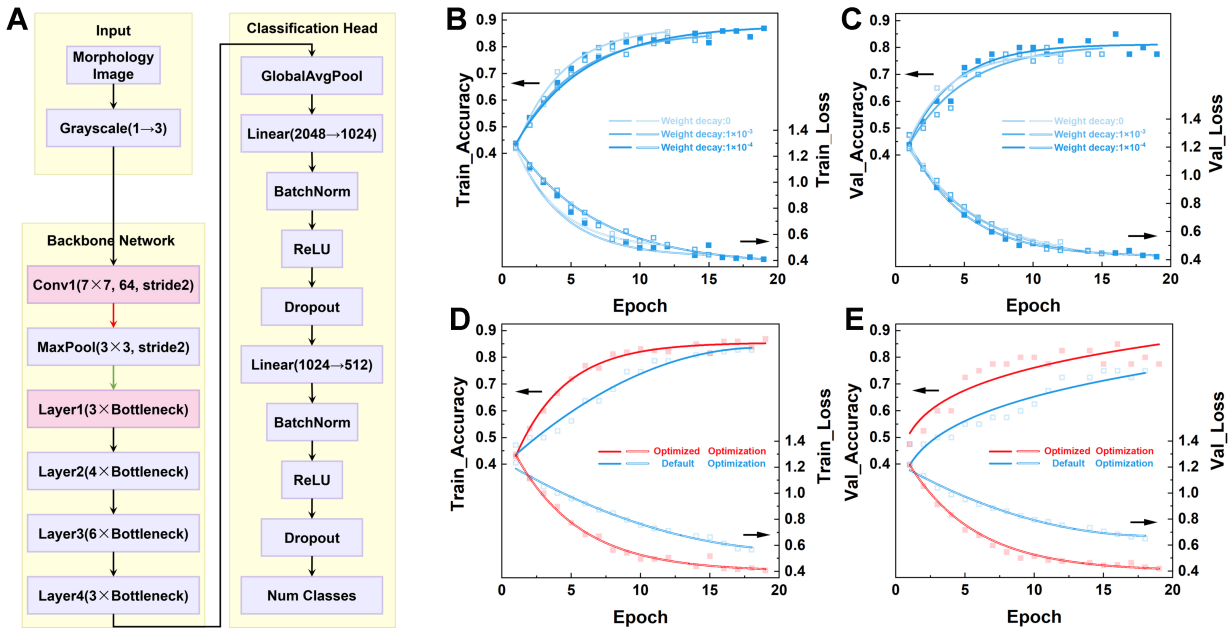


Figure 3. Architectural implementation and classification performance analysis of the gradient-thickness SRO morphology classification model. (A) CNN architecture for SRO morphology classification; (B and C) Accuracy curves and loss function trends for the (B) training set and (C) validation set under different weight decay hyperparameters; (D and E) Comparison of accuracy and loss function trends for the (D) training set and (E) validation set between the default and optimized parameter configurations. SRO: $SrRuO_3$; CNN: convolutional neural network.

features in material images, preserving the shallow general feature extraction capabilities inherited from ImageNet pretraining while enabling the network to learn material-specific texture patterns. Moreover, the original fully connected layers are replaced with an adaptive structure. The reconstructed classification head incorporates predefined parameters: a dimensionality reduction strategy decreases feature channels from 2,048 to 512, followed by batch normalization. The initial dropout rate is set to 0.3 and further optimized in subsequent experiments to mitigate overfitting risks. Residual connections are retained, with improved feature fusion methods. In addition, a squeeze-and-excitation attention mechanism is integrated into the convolutional module, replacing global average pooling with adaptive average pooling. To enhance the performance of the CNN model in SRO morphological classification, we systematically optimized key hyperparameters, including the learning rate, batch size, number of training epochs, optimization algorithm, dropout rate, patience, and weight decay. Specifically, the learning rate was set to 5×10^{-4}, the batch size to 16, the number of training epochs to 30, and the dropout rate to 0.3. The Adam optimizer was used, early stopping patience was set to 3, and weight decay was initially set to 0. To guide parameter selection more effectively, we designed a composite scoring function that simultaneously evaluates validation accuracy, classification balance, and overfitting severity, with respective weights of 50%, 30%, and 20%. This approach significantly improves parameter selection accuracy compared to single-metric evaluation. The formula for this scoring function is as follows[43]:

$$S_{core} = 50\% \times Acc_{val} + 30\% \times F1_{val} - 20\% \times |Loss_{train} - Loss_{val}|$$

Weight decay, a hyperparameter that controls the intensity of L2 regularization, induces model underfitting when set excessively high, while overly small values fail to effectively constrain weight growth. Therefore, three weight decay values (0, 1×10^{-3}, and 1×10^{-4}) were evaluated to optimize model accuracy. The corresponding composite scores were 0.6246, 0.6341, and 0.6348, respectively. The accuracy and loss curves for both the training and validation sets under different weight decay settings are shown in Figure 3B and C.

When the weight decay was set to 0, the validation accuracy stabilized around 77.50%, with the minimum validation loss remaining near 0.5269. Notably, a 10.00% gap between training and validation accuracy was observed, indicating insufficient feature learning due to the lack of regularization, as evidenced by training stagnation and validation loss saturation. When the weight decay was increased to 1×10^{-4}, the model achieved the highest validation accuracy (80.00%) and the lowest validation loss (0.4202). This result suggests that a small but non-zero weight decay effectively mitigates overfitting while enhancing generalization. By comparing the composite scores and curve characteristics, a weight decay of 1×10^{-4} was selected as the optimal parameter configuration, achieving a balance between model performance and robustness. Additionally, different learning rates (5×10^{-4}, 5×10^{-5}, and 1×10^{-5}), batch sizes (8, 16, and 32), and optimizers (Adam, AdamW, and AMSGrad) were systematically evaluated to further refine the performance of the CNN model. The corresponding accuracy and loss curves are presented in Supplementary Figures 1-3. Through a comprehensive grid search analysis considering both quantitative scoring and convergence patterns, the optimal parameter configuration was identified as follows: a learning rate of 5×10^{-4}, batch size of 8, AMSGrad optimizer, and weight decay of 1×10^{-4}. To validate the effectiveness of this optimization approach, comparative evaluations were conducted on the SRO morphological image dataset using both the baseline model (with default parameters) and the optimized model. The results, shown in Figures 3D and E, demonstrate that the optimized model achieves significantly higher accuracy and lower loss values compared to its baseline counterpart. This improvement confirms that the parameter optimization strategy effectively enhances learning capability and generalization performance.

The morphological classification results of the CNN model on the independent test set are illustrated in Figure 4A. The predictions demonstrate high agreement with the ground truth labels for most samples, with output confidence values exceeding 0.86, thereby confirming the model's strong discriminative capability and reliable decision making for typical specimens. However, a low-confidence case (0.57) was observed in the test set, which resulted from surface contaminants or particle adhesion during AFM characterization. These contaminants caused significant topographical interference during probe scanning, ultimately producing AFM images with substantially compromised clarity. To address these inevitable artifacts that may cause low-confidence predictions, we exclusively selected AFM morphological images with confidence scores exceeding 0.80 for the independent test set used in subsequent classification accuracy evaluation, thereby ensuring the reliability of the final classification accuracy. The comprehensive prediction analysis via the confusion matrix (CM) across the entire test set (n = 60) yielded an overall accuracy of 86.67% (52/60), demonstrating the model's robust global classification proficiency, as shown in Figure 4B. Notably, the stripes morphology class dominates the correct predictions, contributing 51.92% of the total accurate classifications. For the particles class (n = 12), balanced recall and precision rates of 75.00% (9/12) are achieved, with 16.67% (2/12) misclassified as stripes and 8.33% (1/12) as smooth. The smooth class (n = 8) exhibits perfect recall (100.00%, 8/8) yet suboptimal precision (72.73%, 8/11), suggesting a tendency toward overprediction. Strikingly, the stripes class (n = 31) achieves superior performance with 87.10% recall (27/31) and 93.10% precision (27/29), highlighting its effectiveness in identifying the majority class. Nevertheless, some misclassifications persist: 9.68% (3/31) are erroneously assigned to particles, while 3.23% (1/31) are mislabeled as smooth. Intriguingly, the stripes-to-particles misclassification rate (9.68%) markedly exceeds the reciprocal particles-to-stripes error (16.67%). The step-like morphology (n = 9) shows the best performance, attaining near-perfect metrics with 88.89% recall (8/9) and 100.00% precision (8/8), with only one instance misclassified as smooth. The receiver operating characteristic (ROC) curves of the model are presented in Figure 4C, with area under the curve (AUC) values for the four morphological categories as follows: smooth (AUC = 0.98), step-like (AUC = 0.98), particles (AUC = 0.89), and stripes (AUC = 0.92). These results corroborate the CM analysis, where the elevated AUC values for the smooth and step-like classes indicate the model's capability to maintain high true positive rates (TPR ≥ 0.9) while achieving low false positive rates (FPR ≤ 0.05). Collectively, the integrated evidence from both CM and ROC metrics

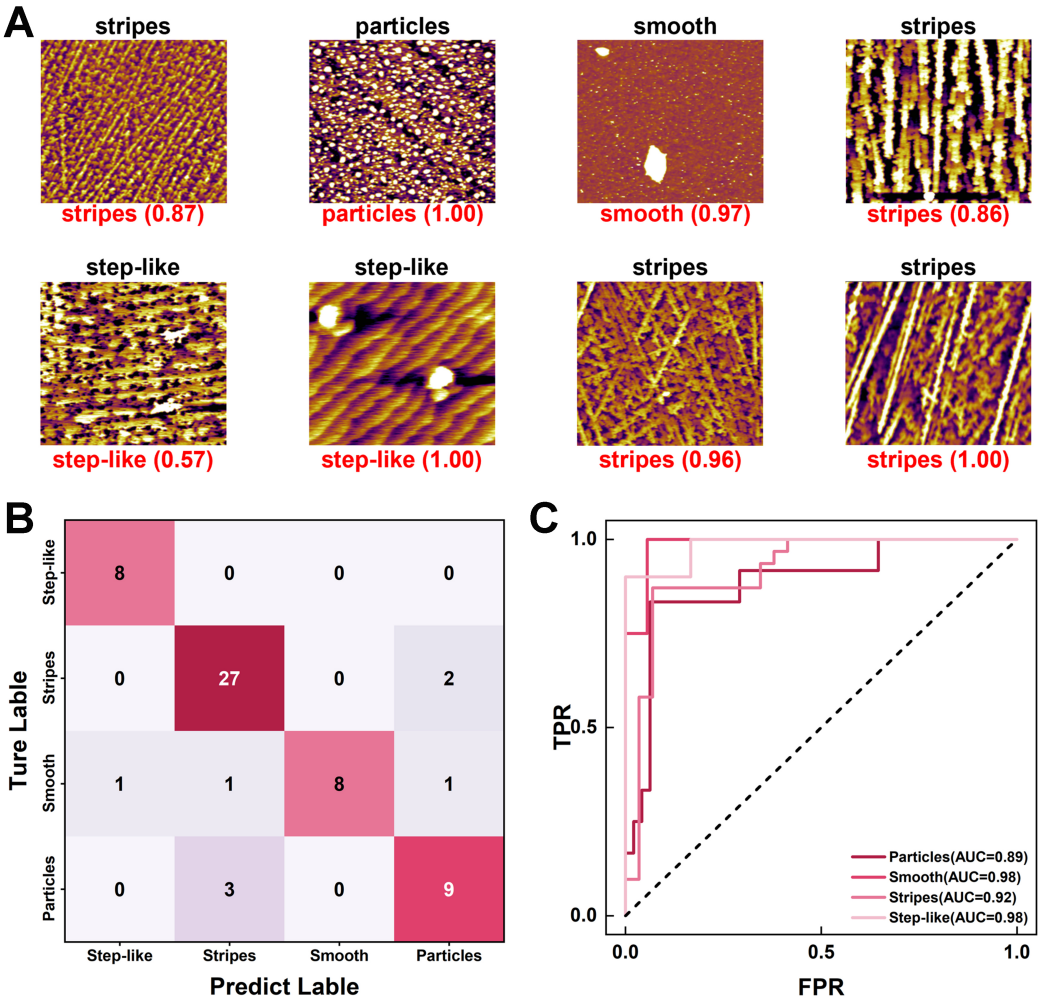


Figure 4. Systematic performance evaluation of the SRO morphological classification model. (A) Visualization of representative classification results from the independent test set; (B) Multiclass CM for classification accuracy assessment; (C) ROC curves with corresponding AUC values across four morphological categories. SRO: $SrRuO_3$; CM: confusion matrix; ROC: receiver operating characteristic; AUC: area under the curve; TPR: true positive rates; FPR: false positive rates.

confirms that the proposed CNN model demonstrates high-precision classification across all morphological types.

CONCLUSIONS

In summary, this study presents an ML framework for the quantitative analysis of morphological features in AFM images of SRO thin films. Systematic optimization, based on a comparative analysis of ResNet architectures and layer freezing strategies, coupled with grid search hyperparameter tuning, yielded an optimal CNN configuration. The implemented architecture achieved 86.67% classification accuracy on an independent test set when categorizing AFM morphological features, including step-like or stripes morphologies. The proposed CNN model demonstrates high accuracy in AFM image classification, as evidenced by experimental validation. This deep learning approach significantly reduces the time and labor costs associated with AFM image quality analysis and shows potential for extension to other material characterization applications.

DECLARATIONS

Authors' contributions

Conceived and coordinated the project: Chen, Q.; Wang, X.; Zhong, G.
Fabricated the substrate and films: Wang, X.; Chen, Q.; Dai, L. (under the guidance of Zhong, X.; Wang, J.; Zhong, G.)
Performed structural analysis and interpretation: Zou, J.; Song, H.
Wrote the manuscript: Wang, X.; Chen, Q.; Zhong, X.; Zou, J.; Zhong, G.
All authors participated in discussions and data analysis.

Availability of data and materials

The original contributions presented in this study are included in the Supplementary Materials. Further inquiries can be directed to the corresponding authors.

Financial support and sponsorship

This work was supported by the National Key Research and Development Program of China (2022YFF0706100), the National Natural Science Foundation of China (Nos. 62474186, 12275230, and 52472139), and the Shenzhen Science and Technology Program (JCYJ20240813155702004).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

Supplementary Materials

Supplementary Materials

REFERENCES

1. Choo, S.; Varshney, S.; Liu, H.; Sharma, S.; James, R. D.; Jalan, B. From oxide epitaxy to freestanding membranes: opportunities and challenges. *Sci. Adv.* **2024**, *10*, eadq8561. DOI PubMed PMC
2. Wang, S.; Li, W.; Deng, C.; et al. Giant electric field-induced second harmonic generation in polar skyrmions. *Nat. Commun.* **2024**, *15*, 1374. DOI PubMed PMC
3. Yang, X.; Han, L.; Ning, H.; et al. Ultralow-pressure-driven polarization switching in ferroelectric membranes. *Nat. Commun.* **2024**, *15*, 9281. DOI PubMed PMC
4. Dou, L.; Yang, B.; Ye, X.; et al. Flexible high-entropy functional ceramics. *Nat. Commun.* **2025**, *16*, 5915. DOI PubMed PMC
5. Liu, K.; Jin, F.; Zhu, T.; et al. Interface-controlled uniaxial in-plane ferroelectricity in $Hf_{0.5}Zr_{0.5}O_2$(100) epitaxial thin films. *Nat. Commun.* **2025**, *16*, 7385. DOI PubMed PMC
6. Ma, C.; Luo, Z.; Huang, W.; et al. Sub-nanosecond memristor based on ferroelectric tunnel junction. *Nat. Commun.* **2020**, *11*, 1439. DOI PubMed PMC
7. Liu, C.; Si, Y.; Zhang, H.; et al. Low voltage-driven high-performance thermal switching in antiferroelectric $PbZrO_3$ thin films. *Science* **2023**, *382*, 1265-9. DOI PubMed
8. Yang, B.; Zhang, Q.; Huang, H.; et al. Engineering relaxors by entropy for high energy storage performance. *Nat. Energy.* **2023**, *8*, 956-64. DOI
9. Guo, Y.; Peng, B.; Lu, G.; et al. Remarkable flexibility in freestanding single-crystalline antiferroelectric $PbZrO_3$ membranes. *Nat. Commun.* **2024**, *15*, 4414. DOI PubMed PMC
10. Qian, J.; Liu, Y.; He, L.; et al. Topological bubble domain engineering for high strain response. *Sci. Adv.* **2025**, *11*, eadw8840. DOI PubMed

11. Ren, Z.; Deng, S.; Shao, J.; et al. Ultrahigh-power-density flexible piezoelectric energy harvester based on freestanding ferroelectric oxide thin films. *Nat. Commun.* **2025**, *16*, 3192. DOI PubMed PMC
12. Yao, Y.; Liu, H.; Hu, Y.; et al. Fluctuating local polarization: a generic fingerprint for enhanced piezoelectricity in Pb-based and Pb-free perovskite ferroelectrics. *Nat. Commun.* **2025**, *16*, 7442. DOI PubMed PMC
13. Tao, A.; Jiang, Y.; Chen, S.; et al. Ferroelectric polarization and magnetic structure at domain walls in a multiferroic film. *Nat. Commun.* **2024**, *15*, 6099. DOI PubMed PMC
14. Pan, Q.; Xiong, Y. A.; Sha, T. T.; et al. Strain-induced tunable enhancement of piezoelectricity in a novel molecular multiferroic material. *Adv. Mater.* **2025**, *37*, e2410585. DOI PubMed
15. Trier, F.; Noël, P.; Kim, J.; Attané, J.; Vila, L.; Bibes, M. Oxide spin-orbitronics: spin–charge interconversion and topological spin textures. *Nat. Rev. Mater.* **2022**, *7*, 258-74. DOI
16. Gao, W.; Zhu, Y.; Wang, Y.; Yuan, G.; Liu, J. A review of flexible perovskite oxide ferroelectric films and their application. *J. Materiomics.* **2020**, *6*, 1-16. DOI
17. Xu, R.; Liu, S.; Saremi, S.; et al. Kinetic control of tunable multi-state switching in ferroelectric thin films. *Nat. Commun.* **2019**, *10*, 1282. DOI PubMed PMC
18. Fan, Y.; Zhang, S.; Xue, Z.; et al. Hidden structural phase transition assisted ferroelectric domain orientation engineering in $Hf_{0.5}Zr_{0.5}O_2$ films. *Nat. Commun.* **2025**, *16*, 4232. DOI PubMed PMC
19. Hafner, J.; Benaglia, S.; Richheimer, F.; et al. Multi-scale characterisation of a ferroelectric polymer reveals the emergence of a morphological phase transition driven by temperature. *Nat. Commun.* **2021**, *12*, 152. DOI PubMed PMC
20. Yao, S.; Jiang, H.; Wen, J.; et al. Electrically modulated photothermal force microscopy for revealing molecular conformation changes during polarization switching at the nanoscale. *Nat. Commun.* **2025**, *16*, 6680. DOI PubMed PMC
21. Cai, J.; Chu, X.; Xu, K.; Li, H.; Wei, J. Machine learning-driven new material discovery. *Nanoscale. Adv.* **2020**, *2*, 3115-30. DOI PubMed PMC
22. Jiang, Y. F.; Peng, H. Y.; Cai, Y.; et al. Adaptive ferroelectric memristors with high-throughput $BaTiO_3$ thin films for neuromorphic computing. *Mater. Horiz.* **2025**, *12*, 6928-37. DOI PubMed
23. Koinuma, H.; Takeuchi, I. Combinatorial solid-state chemistry of inorganic materials. *Nat. Mater.* **2004**, *3*, 429-38. DOI PubMed
24. Barber, Z. H.; Blamire, M. G. High throughput thin film materials science. *Mater. Sci. Technol.* **2008**, *24*, 757-70. DOI
25. Yuan, J.; Stanev, V.; Gao, C.; Takeuchi, I.; Jin, K. Recent advances in high-throughput superconductivity research. *Supercond. Sci. Technol.* **2019**, *32*, 123001. DOI
26. Wang, Z.; Sun, Z.; Yin, H.; et al. Data-driven materials innovation and applications. *Adv. Mater.* **2022**, *34*, e2104113. DOI PubMed
27. Green, M. L.; Choi, C. L.; Hattrick-Simpers, J. R.; et al. Fulfilling the promise of the materials genome initiative with high-throughput experimental methodologies. *Appl. Phys. Rev.* **2017**, *4*, 011105. DOI
28. Zhang, Y.; Tan, Y.; Dong, Y.; et al. High-throughput scanning second-harmonic-generation microscopy for polar materials. *Adv. Mater.* **2023**, *35*, e2300348. DOI PubMed
29. Tang, M.; Dai, L.; Cheng, M.; et al. High-throughput screening thickness-dependent resistive switching in $SrTiO_3$ thin films for robust electronic synapse. *Adv. Funct. Mater.* **2023**, *33*, 2213874. DOI
30. Fang, H.; Wang, J.; Nie, F.; et al. Giant electroresistance in ferroelectric tunnel junctions via high-throughput designs: toward high-performance neuromorphic computing. *ACS. Appl. Mater. Interfaces.* **2024**, *16*, 1015-24. DOI PubMed
31. Zhang, D.; Harmon, K. J.; Zachman, M. J.; et al. High-throughput combinatorial approach expedites the synthesis of a lead-free relaxor ferroelectric system. *InfoMat* **2024**, *6*, e12561. DOI
32. Xiao, R.; Zhang, Y.; Li, M. Automated high-throughput atomic force microscopy single-cell nanomechanical assay enabled by deep learning-based optical image recognition. *Nano. Lett.* **2024**, *24*, 12323-32. DOI PubMed
33. Thomas-Chemin, O.; Janel, S.; Boumehdi, Z.; et al. Advancing high-throughput cellular atomic force microscopy with automation and artificial intelligence. *ACS. Nano.* **2025**, *19*, 5045-62. DOI PubMed
34. Gregoire, J. M.; Van Campen, D. G.; Miller, C. E.; Jones, R. J.; Suram, S. K.; Mehta, A. High-throughput synchrotron X-ray diffraction for combinatorial phase mapping. *J. Synchrotron. Radiat.* **2014**, *21*, 1262-8. DOI PubMed
35. Maruyama, S.; Ouchi, K.; Koganezawa, T.; Matsumoto, Y. High-throughput and autonomous grazing incidence X-ray diffraction mapping of organic combinatorial thin-film library driven by machine learning. *ACS. Comb. Sci.* **2020**, *22*, 348-55. DOI PubMed
36. Ludwig, A. Discovery of new materials using combinatorial synthesis and high-throughput characterization of thin-film materials libraries combined with computational methods. *npj. Comput. Mater.* **2019**, *5*, 205. DOI
37. Kimmig, J.; Zechel, S.; Schubert, U. S. Digital transformation in materials science: a paradigm change in material's development. *Adv. Mater.* **2021**, *33*, e2004940. DOI PubMed PMC

38. Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547-55. DOI PubMed
39. Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse molecular design using machine learning: generative models for matter engineering. *Science* **2018**, *361*, 360-5. DOI PubMed
40. Himanen, L.; Geurts, A.; Foster, A. S.; Rinke, P. Data-driven materials science: status, challenges, and perspectives. *Adv. Sci.* **2019**, *6*, 1900808. DOI PubMed PMC
41. Hardian, R.; Liang, Z.; Zhang, X.; Szekely, G. Artificial intelligence: the silver bullet for sustainable materials development. *Green. Chem.* **2020**, *22*, 7521-8. DOI
42. Xu, X.; Ma, W.; Yan, B. An electrodeposited nano-porous and neural network-like Ln@HOF film for SO_2 gas quantitative detection *via* fluorescent sensing and machine learning. *J. Mater. Chem. A.* **2021**, *9*, 26391-400. DOI
43. Miseta, T.; Fodor, A.; Vathy-Fogarassy, Á. Surpassing early stopping: a novel correlation-based stopping criterion for neural networks. *Neurocomputing* **2024**, *567*, 127028. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.


© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.