

Article

Accelerated Design for Perovskite-Oxide-Based Photocatalysts Using Machine Learning Techniques

Xiuyun Zhai ^{1,*}  and Mingtong Chen ^{2,*}¹ College of Intelligent Manufacturing, Hunan University of Science and Engineering, Yongzhou 425199, China² Public Experimental Teaching Center, Panzhihua University, Panzhihua 617000, China

* Correspondence: cqfbb2008@shu.edu.cn (X.Z.); cmt2711@126.com (M.C.)

Abstract: The rapid discovery of photocatalysts with desired performance among tens of thousands of potential perovskites represents a significant advancement. To expedite the design of perovskite-oxide-based photocatalysts, we developed a model of ABO₃-type perovskites using machine learning methods based on atomic and experimental parameters. This model can be used to predict specific surface area (SSA), a key parameter closely associated with photocatalytic activity. The model construction involved several steps, including data collection, feature selection, model construction, web-service development, virtual screening and mechanism elucidation. Statistical analysis revealed that the support vector regression model achieved a correlation coefficient of 0.9462 for the training set and 0.8786 for the leave-one-out cross-validation. The potential perovskites with higher SSA than the highest SSA observed in the existing dataset were identified using the model and our computation platform. We also developed a webserver of the model, freely accessible to users. The methodologies outlined in this study not only facilitate the discovery of new perovskites but also enable exploration of the correlations between the perovskite properties and the physicochemical features. These findings provide valuable insights for further research and applications of perovskites using machine learning techniques.

Keywords: photocatalyst; machine learning; ABO₃-type perovskites; specific surface area; prediction

**Citation:** Zhai, X.; Chen, M.Accelerated Design for Perovskite-Oxide-Based Photocatalysts Using Machine Learning Techniques. *Materials* **2024**, *17*, 3026. <https://doi.org/10.3390/ma17123026>

Academic Editor: Fernando Gomes

Received: 22 March 2024

Revised: 5 June 2024

Accepted: 8 June 2024

Published: 20 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Perovskite oxides, with the chemical formula ABO₃, are a broad category of compounds with similar structures [1–4]. The earliest compound in this group of compounds is calcium titanate (CaTiO₃), which was discovered in calcium titanium ore [5]. In perovskite materials, the B-site typically consists of cations with smaller ionic radii. These cations co-ordinate with oxygen anions and occupy the corners of the octahedral structure in the cubic lattice [4,6]. The typical structure of ABO₃-type perovskites has Pm3m symmetry (Figure 1). Perovskite structures offer significant flexibility for exploration due to the potential replacement of ions at various positions with those of other elements or groups with similar radii [7,8]. This versatility results in a diverse array of perovskite compounds, offering a vast potential for specific applications [9,10]. Considering element doping, the potential number of perovskite compounds can reach approximately 10⁷ [11]. These perovskites exhibit a remarkable diversity, offering several superior properties, including high light absorption coefficients [12,13], suitable and adjustable bandgap [14], narrow emission bandwidth [15], long carrier diffusion lengths [16,17], high carrier mobilities [18,19], and cost-effective solution processability [20]. These attributes make perovskites highly promising materials in various fields such as solar cells [21,22], photocatalytic materials [7,23], and photodetectors [24].

Conventional methods of material development typically rely on repeated trial-and-error processes until the desired material properties are achieved [25,26]. Although first principle calculation of materials can provide key characteristics of materials without the

need for experimental synthesis [27,28], it involves extensive equation calculations and approximations. However, both experimental methods and first principles calculations face challenges in meeting the demand for large-scale, rapid, and efficient prediction within the vast chemical search space and complex crystal component structures of perovskite-type materials.

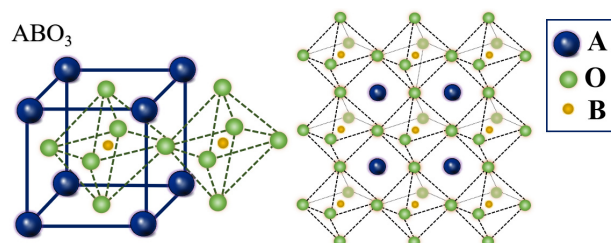


Figure 1. A schematic illustration of the crystal cell of a typical ABO_3 -type perovskite.

Machine learning (ML), as crucial branch of artificial intelligence, can rapidly and effectively evaluate or predict research objects [29–32]. This highlights the significant potential of ML in the design, synthesis, physical property exploration, and application research of perovskite-type materials [33,34]. ML has been successfully used to accelerate the development of perovskite-oxide-based photocatalysts and solar cell materials in recent years [35–37]. ML applications in ABO_3 -type perovskites primarily include predictions related to stability, bandgap, crystal structure, and formability (Table S1 of the Supporting Information). As Experimental data on perovskite synthesis continue to grow, ML is anticipated to play an increasingly vital role in the advancement of perovskite-type photocatalysts [38,39].

ABO_3 -type perovskite compounds are a novel class of semiconductor photocatalysts [40], where specific surface area (SSA) is an important indicator of photocatalytic activity [41]. A larger SSA enhances the availability of active sites, thereby improving photocatalytic performance [42]. In this study, a model of ABO_3 -type perovskites for designing perovskite-type photocatalysts with desired SSA is proposed. The model was developed using atomic and experimental parameters with ML methods. The method involved several steps, including data collection, feature engineering, model construction, web-service development, virtual screening, and mechanism elucidation.

ABO_3 -type perovskites are highly promising semiconductor photocatalysts, drawing increasing interest from researchers [43,44]. This study focused on rapidly screening promising ABO_3 -type perovskites for photocatalytic applications and accelerating the design of perovskite-oxide-based photocatalysts. The main contributions of our work are outlined as follows: (1) we compiled a dataset of ABO_3 -type perovskites by gathering experimental data from published references. (2) We developed a support vector regression (SVR) model for synthesizing ABO_3 -type perovskites using the sol–gel method. The model demonstrates high accuracy and good generalization, with correlation coefficients (R) of 0.9462 for the training set and 0.8786 for leave-one-out cross-validation (LOOCV). (3) Key factors influencing the SSA of ABO_3 -type perovskites were identified using forward and backward selection methods based on the SVR model with radial basis function (RBF). (4) The model identified promising ABO_3 -type perovskites with potentially high SSA. (5) An online web service was developed to facilitate rapid and effective prediction of SSA for ABO_3 -type perovskites. This service is accessible at http://1.14.49.110/online_predict/Perovskite (accessed on 7 June 2024).

The other sections of this paper are organized as follows. In Section 2, the main steps involved in constructing the SSA model of ABO_3 -type perovskites are presented in detail, along with the computational details. Feature selection, model construction, model application, and mechanism exploration are presented in Section 3. Section 4 presents the conclusions of the study.

2. Material and Methods

2.1. Perovskite Model Framework

The SSA model for ABO₃-type perovskites was developed through six primary steps, as shown in Figure 2. These steps include data collection, feature engineering, model construction, web-service development, virtual screening, and mechanism mining. The details of each step are as follows.

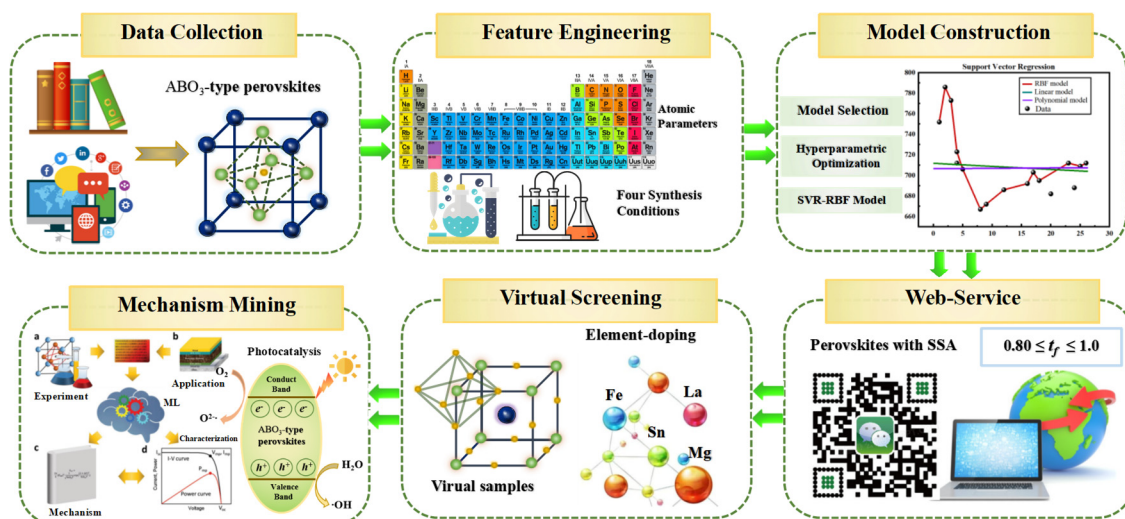


Figure 2. An illustration of the six steps involved in the construction of the model.

2.1.1. Data Collection

Data quality is crucial for ML model quality and fundamentally determines its reliability [45]. In most ML research, the upper limit of the model performance depends on the quality of data and features selected, whereas the choice of models and algorithms can only help approach this upper limit infinitely [46]. In this study, we established a reliable dataset by collecting extensive data of ABO₃-type perovskites. Specifically, we compiled data from 99 perovskite samples sourced from the experimental results reported in the previous literature (Table S2). These ABO₃-type perovskites were synthesized using the sol–gel technique [47], which involves transforming inorganic compounds or metal alkoxides into oxides or other solid compounds through solidification process involving solution, sol, gel, and subsequent heat-treat treatment. The three values (namely 1, 2, and 3) in the *PM* (synthetic mode) column in Table S2 represent three modes of the sol–gel technique: traditional, auto-combustion, and citrate methods. The target variable of the dataset, SSA, denotes the surface area per unit mass of perovskite crystals, measured in square meters per gram (m²g^{−1}).

The dataset comprised 25 features as the inputs of the ML model (Table S3), including the 4 synthesis conditions (calcination temperature (*CT*), calcination time (*AH*), drying temperature (*DT*), and *PM*) and 21 atomic parameters. The atomic parameters were derived using the molecular formula of the samples through the Online Computational Platform of Material Data Mining (<http://materials-data-mining.com/ocpmdm/>) (OCPMDM) [48,49]. The dataset was divided into two parts: the training set comprising 85 samples for constructing the model and the testing set comprising 14 samples (marked with asterisks in Table S2) for external validation. The following rules were observed when dividing the dataset:

- (1) The SSA values of the samples in the testing set were kept within the range of values observed in the training set to prevent out of range predictions.
- (2) Samples in the dataset were sorted based on the SSA values. Samples in the testing set were selected at regular intervals from this sorted dataset to avoid biased predictions.

In addition, a validation set was established to further verify the model. The validation set comprised three samples (marked with hashes in Table S2) obtained from recently published papers.

2.1.2. Feature Engineering

Ensuring the stability and generalization of the prediction model ensures managing high linear correlations between features used in modeling. An important step in this process involves removing features with strong correlations to build an effective model. To achieve this, Pearson correlation coefficients between the 25 features were calculated, and the correlation matrix is presented in Figure S1. Maximum-relevance minimum-redundancy (mRMR) scores of all features were calculated (Figure S2). Features with lower mRMR scores were eliminated from the dataset if their correlation coefficient with another feature was greater than 0.95. Twenty-three features were retained through this correlative analysis. Features such as the ratio of ionic radius (R_A/R_B) and unit cell lattice edge (a) were eliminated due to their correlation with other features. A feature selection method based on SVR was used to identify the most representative subset from the feature pool. This approach ensures that the selected subset contains essential information with minimal redundancy.

2.1.3. Model Construction

The ML model was designed to establish the relationship between SSA of ABO₃-type perovskites and the corresponding features. An SVR model with RBF kernel was chosen to accelerate the design of perovskite-type materials. The hyperparameters of the model were optimized using the grid search technique and cross-validation to enhance the prediction performance of the model. The external testing set comprising 14 samples, LOOCV, and an independent validation set were utilized to test the reliability of the model.

2.1.4. Webserver Development

A webserver refers to software modules operating over a network. It is service-oriented based on distributed programs, allowing users to access data from different web terminals across different locations. Implementing webservers based on the model simplifies the prediction tasks for the clients. Scientists researching perovskite-type materials can use online predictions through the webserver without mastering the intricacies of the ML model.

2.1.5. Virtual Screening

The general molecular formula of ABO₃-type perovskites is A₁_y(A₂_nA₃_(1-y-n))B₁_x(B₂_mB₃_(1-x-m))O₃, where x and y range from 0.6 to 1.0 in increments of 0.01, and $0 \leq n \leq 0.4$, $0 \leq m \leq 0.4$. This formula facilitates the creation of millions of ABO₃-type perovskites. It is generally assumed that stable perovskites can form if the tolerance factor (t_f) falls within the range of 0.8~1.0. The formula for t_f is as follows:

$$t_f = \frac{R_A + R_O}{\sqrt{2}(R_B + R_O)} \quad (1)$$

where R_A , R_B , and R_O represent the ionic radii of A-site, B-site, and O, respectively.

The virtual samples were generated to screen perovskites with high SSA considering the elements present in the perovskites from the existing dataset and the conditions outlined above.

2.1.6. Mechanism Mining

Mechanism mining is essential in materials research, particularly for understanding how the features affect the SSA of ABO₃-type perovskites. The features that most significantly impact the SSA of ABO₃-type perovskites can be identified by leveraging the

interpretability of ML results. Mechanism mining can help to adjust the material design approach and provide precise guidance for material synthesis. The effectiveness of mechanism mining directly influences the practical application and usability of the models developed.

2.2. Computational Details

In this study, we conducted most of the ML calculations using HyperMiner software package (2009 edition) [50] and the in-house OCPMDM [51–53]. HyperMiner is available for free download from the website of our laboratory (<http://materials-data-mining.com/home>). Detailed instructions for using OCPMDM can be accessed from the webserver link (<http://materials-data-mining.com/ocpmdm/>).

3. Results and Discussion

To enhance the predictive performance of ML models, it is essential to consider different types of models and compare their predictive accuracy based on the dataset's characteristics and the suitability of the algorithms. The SVR algorithm offers strict mathematical theoretical support, strong interpretability, and robustness. However, its training process requires significant computational resources and storage capacity, making it more suitable for smaller sample sets.

In this study, we used the widely used ML algorithms to establish the predictive models of SSA of ABO₃-type perovskites. These algorithms include decision tree regression (DTR), gradient boosting regression (GBR), partial least squares (PLS), relevance vector machine (RVM), SVR-RBF, SVR with linear kernel function (SVR-LKF), SVR with polynomial kernel function (SVR-PKF), random forest regression (RFR), and back propagation neural network (BPNN). The root mean square error (RMSE) and Pearson correlation coefficient (R) of the LOOCV results of the nine algorithms are presented in Table 1.

Table 1. The LOOCV results of the nine algorithms.

Index	DTR	GBR	PLS	RVM	SVR-RBF	SVR-LKF	SVR-PKF	RFR	BPNN
RMSE	7.570	5.798	8.588	6.304	4.895	8.408	7.286	7.618	10.301
R	0.688	0.817	0.504	0.791	0.870	0.540	0.679	0.640	0.443

The model using the SVR-RBF algorithm was more effective for predicting the SSA of ABO₃-type perovskites compared to the other algorithms. This conclusion is supported by its lowest RMSE and highest Pearson correlation coefficient (R) in the LOOCV results. SVR, which operates on the principle of structural risk minimization, addresses issues such as uncertain network structure, overfitting, underfitting, and local minima commonly encountered in some algorithms such as artificial neural networks [54]. It is widely considered as one of the best methods for small sample regression problems [55]. Leveraging slack variables and kernel functions, SVR is effective at handling situations where the data are linearly inseparable. RBF kernel is the most commonly used kernel function in SVR, known for its applicability to both small and large sample problems, as well as high-dimensional and low-dimensional datasets. This analysis shows that SVR-RBF is the most suitable for the dataset. Therefore, it was selected as the algorithm to construct the predictive model for the SSA of ABO₃-type perovskites.

3.1. Feature Selection and Analysis

During the ML training process, datasets often contain many samples and features, some of which offer little or no value for modeling. When a significant proportion of the dataset comprises irrelevant features, it can prolong the model training time and increase the risk of underfitting [56]. Conversely, if a subset of data with minor impact account for a large proportion of the dataset, it can prolong the model training time and increase the risk of overfitting of the model [57]. Consequently, feature selection is a key step in building a

reliable model [58]. The objective of feature selection is to identify the optimal subset of features from the original feature pool, one with essential information and less redundancy.

Three approaches based on SVR were used to identify the optimal subset. These approaches include forward selection method (FSM) [59], backward selection method (BSM) [60], and genetic algorithm (GA) [61,62]. FSM is a heuristic method that starts with an empty set and gradually adds features to the feature subsets based on importance. At each stage, a feature that maximizes the model performance is selected to train the ML model. This iterative process continues until the main features contributing significantly to the model are identified and retained for modeling. The principle of BSM is similar to FSM, except that all features are initially included in training the models. Subsequently, iteratively, the feature that contributes the least to model performance is removed and the remaining features are used to train the ML model. GA is a global optimization algorithm that simulates the evolutionary principles observed in biological systems, where only the fittest survives. The search processes of the three methods (illustrated in Figure 3) continue until a feature subset closely approaching the optimal solution is obtained. Notably, the RMSE value first decreases with an increase in the feature number and gradually increases after reaching a minimum value (Figure 3a,b). In addition, the RMSE initially decreases sharply with an increase in the iteration number (Figure 3c). The RMSE reaches its minimum value after 17 iterations and then stabilizes. Conversely, the parameter “score” exhibits an opposite trend to RMSE for the three methods (Figure 3). Comparison of the three methods showed that the FSM-SVR method was superior to the other algorithms, achieving the lowest RMSE and highest score. According to the results of the FSM-SVR method, the optimal subset comprised 12 variables: CT , AT , DT , PM , Z_{IA} , Δfus_A , T_{mA} , T_{mB} , T_{bA} , ρ_A , EA_a , and EA_b .

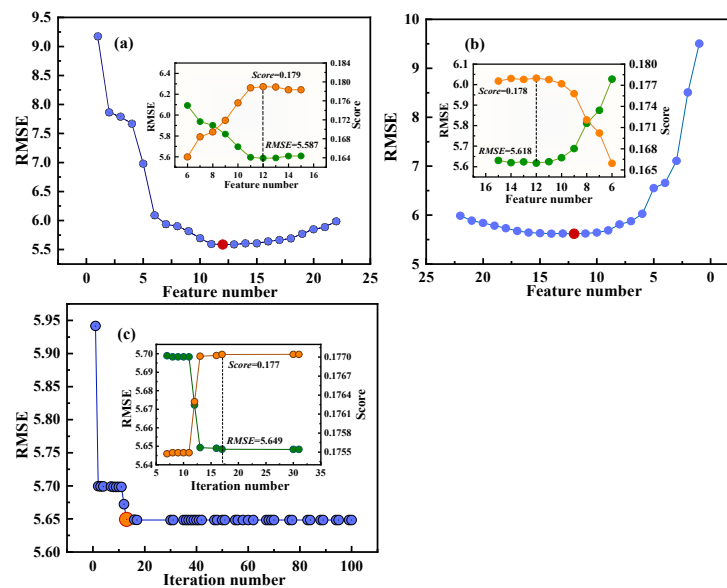


Figure 3. Feature screening using (a) FSM-SVR, (b) BSM-SVR, and (c) GA-SVR. The blue and red circles represent RMSE and the minimum RMSE, respectively. In the illustration, the green and orange circles represent RMSE and Score, respectively.

To further refine the dataset used for modeling, a new dataset was generated incorporating these 12 features and six principal components [63] (P_{CA1} , P_{CA2} , P_{CA3} , P_{CA4} , P_{CA5} , and P_{CA6}) calculated based on the 12 features. The formulas of the principal components are presented as Equations (S1)–(S6) in the Supporting Information. Moreover, nine key features were identified using BSM-SVR method for modeling: CT , DT , PM , T_{mB} , ρ_A , EA_a , EA_b , P_{CA3} , and P_{CA6} .

3.2. Model Construction

To establish an effective model, it is not only essential to select appropriate algorithms and key features for modeling but also to optimize the hyperparameters of the algorithms and evaluate the models using appropriate methods [64].

3.2.1. Optimizing Hyperparameters

Hyperparameter optimization [65] involves finding the optimal combination of hyperparameters to maximize model performance. Therefore, the three hyperparameters [54] (namely, ϵ , C , and γ) of SVR-RBF must be adjusted to enhance accuracy and generalization ability of the SVR-RBF model. The parameter ϵ controls the smoothness of regression curves, influencing the model's tolerance for errors. C is a regularization constant that controls the penalty intensity applied to the model based on errors encountered during training. This parameter determines the trade-off between errors and complexity of the model. The parameter γ is a crucial hyperparameter that controls the kernel function, determining the impacts of sample points on the model. The ranges of ϵ , C , and γ used to optimize the model are (0.01, 0.09), (1, 100), and (0.5, 1.5), and the corresponding step sizes are 0.02, 2, and 0.1, respectively. The optimization process involved using the grid search method with 10-fold cross-validation to find the optimal combination of these hyperparameters. The search process is illustrated in Figure 4. The optimal values of ϵ , C , and γ were 0.07, 97, and 0.9, resulting in a minimal RMSE value of 4.584, respectively.

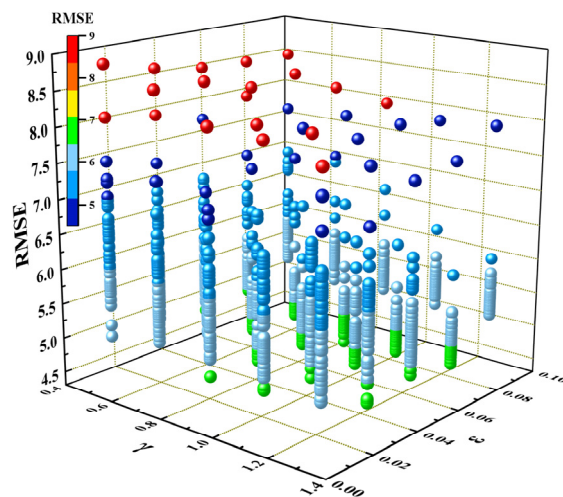


Figure 4. RMSE versus γ and ϵ obtained in the optimization process of the algorithm hyperparameters.

3.2.2. Establishing Model

The SVR-RBF model was constructed using the key features and the optimal hyperparameters. The expression of the model is shown below:

$$SSA = \sum_{i=1}^n \beta_i \exp\left((-0.9) \times ||X - X_i||^2\right) - 0.0927 \quad (2)$$

where n and β_i represent the corresponding number and Lagrange multiplier of the support vectors, respectively. X and X_i denote the unknown vector and the support vector, respectively.

R , the coefficient of determination (R^2), and RMSE were indices to evaluate the performance of the SVR-RBF model by comparing the experimental and predicted SSA values of ABO_3 -type perovskites. Figure 5a illustrates the experimental SSA versus the predicted SSA for the samples in the training set. The SVR-RBF model demonstrated relatively high accuracy, with R , R^2 , and RMSE values of 0.9462, 0.8916, and 3.2557, respectively.

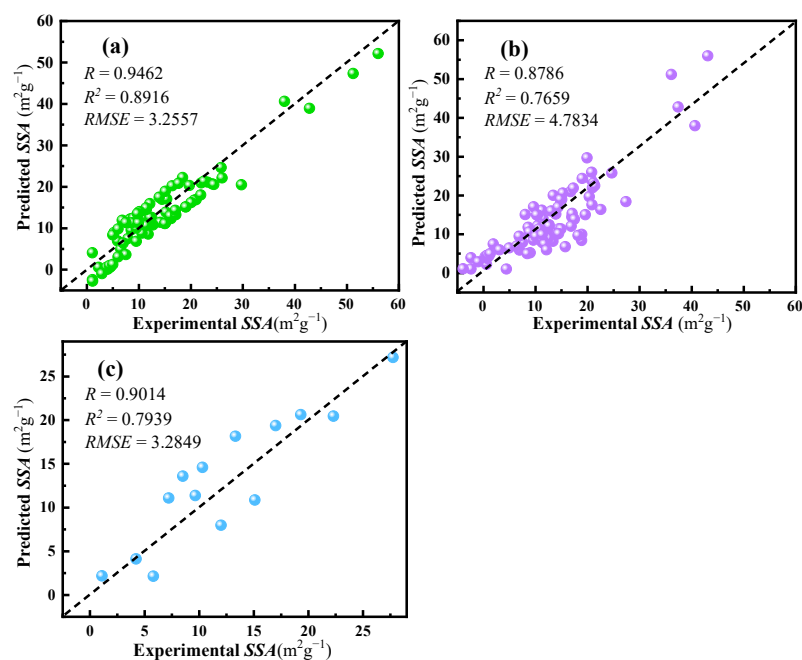


Figure 5. The experimental versus predicted SSA of the samples in (a) the training set; (b) LOOCV; (c) the testing set.

3.2.3. Model Evaluation

LOOCV, the testing set, and the validation set were used to verify the predictive performance of the ML model, with results shown in Figure 5b, Figure 5c, and Table 2, respectively. The model exhibited higher R and lower RMSE for LOOCV and the external set, indicating its effectiveness and reliability. The maximum relative error of the samples was +0.215, with very low relative errors of the other two samples, indicating that the ML model had robust predictive performance (Table 2). The higher relative prediction error observed for sample 101 can be attributed to the absence of the molecular formula of Gd in the samples in the training set, which is the element at the A-site of sample 101. The prediction range of the model can be expanded by collecting more training samples of ABO₃-type perovskites containing Gd or other elements. In addition, the SVR model underwent 100 random 10-fold cross-validations using the training samples. The average R, R², and RMSE values of the 10-fold cross-validations were 0.8695, 0.7538, and 4.8979, respectively. These results further demonstrate the robustness and generalization capability of the SVR-RBF model.

Table 2. The predictive results based on the validation set.

No.	Molecular Formula	Experimental SSA (m ² g ⁻¹)	Predictive SSA (m ² g ⁻¹)	Relative Error
100 [#]	LaFeO ₃	7	6.516	-0.0691
101 [#]	GdCoO ₃	8.69	10.558	+0.215
102 [#]	LaMnO ₃	25	23.830	-0.047

3.3. Model Application

Despite construction of several effective ML models across various fields [66], most of them are not readily accessible to experimental researchers. We adopted two strategies (developing an online web service and conducting virtual screening) to facilitate the utilization of these models by scientists conducting material experiments.

3.3.1. Online Web Service

We established a webserver based on the SVR-RBF model to simplify the prediction of SSA for ABO₃-type perovskites, aiding in designing materials with desired properties. This webserver offers efficiency, convenience, and flexibility in guiding the design and development of ABO₃-type perovskites. A screenshot of the webserver (http://1.14.49.110/online_predict/Perovskite accessed on 7 June 2024) is shown in Figure 6. To use the webserver, users input the chemical formulas of ABO₃-type perovskites into the corresponding dialog box and select values for the four synthetic parameters need be selected from dropdown lists. The guidelines for inputting the molecular formulas of perovskites are provided in small characters at the bottom of Figure 6. The OCPMDM tool automatically generates the atomic parameters based on the formulas provided, eliminating the need for users to enter them manually. Subsequently, the predictive value of SSA of the sample is generated by clicking the “Predict” button. This entire predictive process typically takes a few seconds.

Online Web Prediction Server of Perovskite

- 1 $A1_y(A2_nA3_{(1-y-n)})B1_x(B2_mB3_{(1-x-m)})O_3$
 $0.6 \leq x, y \leq 1; 0 \leq n, m \leq 0.4$
- 2 $t_f = \frac{R_A + R_O}{\sqrt{2}(R_A + R_O)}$ $0.8 \leq t_f \leq 1$
- 3 SSA prediction of ABO₃-type perovskites
- 4 Visual inspection for desired samples

Please input the molecular formula and four synthesis parameters of ABO₃-type perovskite to the corresponding boxes, where CT, AH, DT and PM refer to calcination temperature (°C), calcination time (h), drying temperature (°C) and synthetic mode, respectively. Then, please click the “Predict” button to obtain predictive SSA of the new sample.

Please input perovskite information:

Perovskite Formula	LaMnO3	CT	400	▼	AH	2	▼	DT	80	▼	PM	1	▼
--------------------	--------	----	-----	---	----	---	---	----	----	---	----	---	---

Model Prediction:

Predict

Please note the following points when entering the formulas of ABO₃-type perovskites.

(1) In the formula $A1_y(A2_nA3_{(1-y-n)})B1_x(B2_mB3_{(1-x-m)})O_3$ of ABO₃-type perovskites, the ranges of x and y with the step of 0.01 are $[0.6, 1.0]$, $0 \leq n \leq 0.4$ and $0 \leq m \leq 0.4$.

(2) Elements in A-site:

- ✦ La + [Bi, Sr, Ca] + Ba
- ✦ Ca + Ag + La
- ✦ Na + La
- ✦ [Zn, Pr, Sr, Ga]

(3) Elements in B-site:

- ✦ [Co, Fe, Cu, Ni, Mg, Al] + [Mg, Co, Ni, Cu, Fe, Bi, Ru]
- ✦ [Ti, Cr, Ta, Mn]

Figure 6. A screenshot of the web service for predicting SSA of ABO₃-type perovskites.

3.3.2. Virtual Screening

The results based on the LOOCV and external test indicated the effectiveness and reliability of the SVR-RBF model in predicting the SSA of ABO₃-type perovskites. Consequently, the model was used to predict the SSA of virtual samples to identify potential candidates. A schematic representation of the virtual screening of ABO₃-type perovskites is shown in Figure 7.

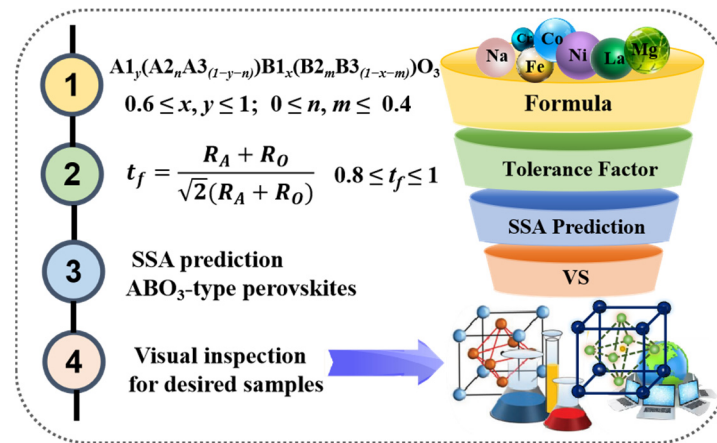


Figure 7. Virtual screening of ABO₃-type perovskites.

The virtual screening process is summarized below:

(1) In the formula $A1_y(A2_nA3_{(1-y-n)})B1_x(B2_mB3_{(1-x-m)})O_3$ for ABO₃-type perovskites, the values of x and y range from 0.6 to 1.0 in increments of 0.01, with $0 \leq n \leq 0.4$ and $0 \leq m \leq 0.4$.

(2) The tolerance factor t_f of perovskites ranged from 0.8 to 1.0 to ensure a stable perovskite structure.

(3) The A-site and B-site allows for a maximum of three and two different doping ions, respectively.

(4) When the first element in the A-site space is La, the second element can be Bi, Sr, or Ca with a doping ratio ranging from 0.0 to 0.4. The third element in the A-site space is Ba with the remaining doping ratio.

(5) When the first element in the A-site space is Ca, the second element in A-site space is Ag and the third element is La with the remaining doping ratio.

(6) When the first element in the A-site space is Na, the second element is La with the remainder of the doping ratio.

(7) The elements in A-site space can be Zn, Pr, Sr, or Ga without doping other elements.

(8) When the first element in the B-site is Co, Fe, Cu, Ni, Mg, or Al, the second element can be Mg, Co, Ni, Cu, Fe, Bi, or Ru, with the remaining doping ratio, ensuring that the second element is different from the first element.

(9) The element in the B-site can be Ti, Cr, Ta, or Mn without doping other elements.

(10) The ranges of CT , AT , and DT are (400 °C, 1000 °C), (2h, 15h), and (80 °C, 300 °C), respectively, with steps of 50 °C, 1 h, and 10 °C, respectively.

(11) The values of PM can be 1, 2, or 3.

We identified the samples with potentially higher SSA than the highest SSA (56 m²g⁻¹) observed in the existing dataset through virtual screening and predictive calculations using the SVR-RBF model. The top two candidates are presented in Table 3, with the highest predictive SSA value reaching 67.884 m²g⁻¹.

Table 3. The top two candidates with potentially higher SSA than the dataset value.

No.	Molecular Formula	SSA (m ² g ⁻¹)	DT (°C)	AH (h)	PM	Fisher (1)	Fisher (2)
1	La _{0.61} Ba _{0.39} TiO ₃	67.884	280	9	1	0.953	1.511
2	La _{0.51} Ba _{0.49} TiO ₃	66.158	280	9	1	0.813	1.471

3.4. Mechanism Mining

We evaluated the relationships between the key features of the model and SSA to further explore the synthesis mechanism of ABO₃-type perovskites. The optimal conditions for the samples with higher SSA were identified by pattern recognition.

3.4.1. Relationships between Key Features and SSA

The relationships between the key features and SSA are explored, and the findings are presented in Figure 8. The results indicate that ABO₃-type perovskites may exhibit higher SSA when the drying temperature is either below 120 °C or above 250 °C (Figure 8a). Excessive calcination time tends to result in low SSA. In addition, the excessive melting point of B-site was not conducive to achieving a high SSA (Figure 8b). Conversely, a higher density of the A-site correlated with a higher SSA. Furthermore, a higher P_{CA3} value and a lower P_{CA6} value were associated with higher SSA (Figure 8c). A larger electron affinity at the A-site and a B-site electron affinity ranging between 0 and 40 kJ/mol potentially generated higher SSA (Figure 8d).

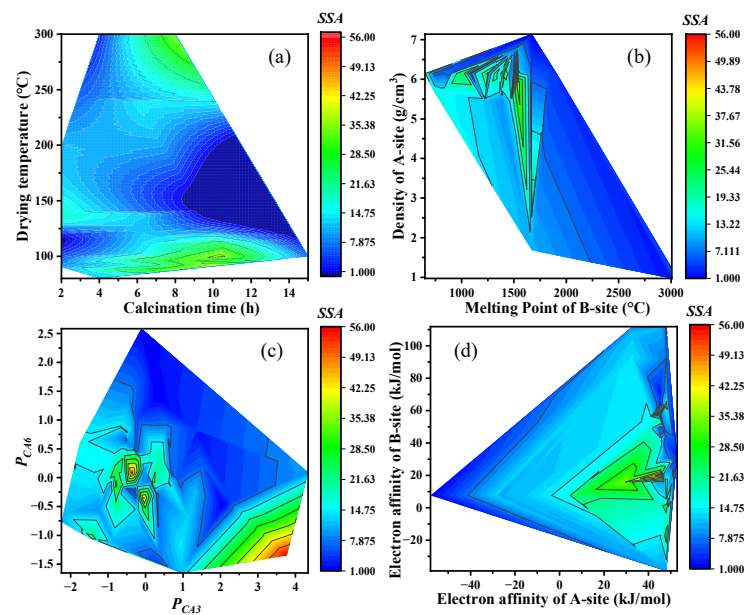


Figure 8. The relationships between SSA and the key features. (a) Calcination time and Drying temperature; (b) Melting point of B-site and Density of A-site; (c) P_{CA3} and P_{CA6} ; (d) Electron affinity of A-site and Electron affinity of B-site.

3.4.2. Pattern Recognition

The samples were divided into two groups based on their SSA: “superior” with higher SSA and “inferior” with lower SSA. The samples frequently have relatively high SSA in the superior region of the projection diagram. The projection of pattern recognition was obtained through the Fisher discriminant analysis method [67,68] and is illustrated in Figure 9. A clear distribution pattern was observed for the two categories of samples (Figure 9). The results showed that most of the superior samples were clustered in a rectangular area of the projection graph, accounting for approximately 90.24% of the total samples in that region. This indicates a significant concentration of samples with higher SSA in a specific area of the projection. The results of pattern recognition for all samples in the training set are summarized in Table 4. The two categories of samples exhibited F1_score equal to or exceeding 0.8, indicating robust performance of the pattern recognition model.

The top two samples with potentially higher SSA are highlighted by purple rectangles in Figure 9. The samples should fall within this rectangular region to achieve new ABO₃-type perovskites with optimal SSA. The expressions defining the superior region are as follows:

$$-0.4618 \leq \text{Fisher}(1) \leq 1.565 \quad (3)$$

$$-0.0858 \leq \text{Fisher}(2) \leq 1.587 \quad (4)$$

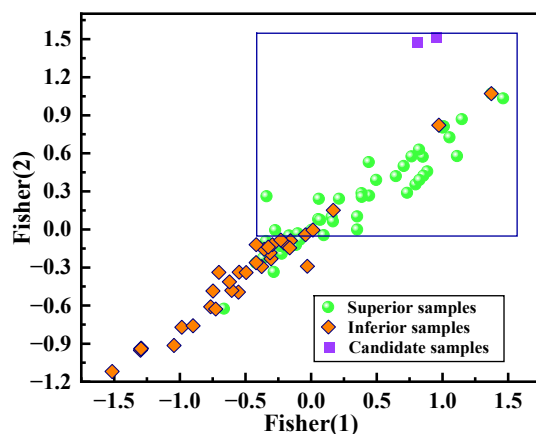


Figure 9. The projection of pattern recognition.

Table 4. The evaluation results of pattern recognition.

Index	“Superior”	“Inferior”
True Positives	37	32
False Positives	4	12
False Negatives	12	4
Precision	0.902	0.727
Recall	0.755	0.889
F1_score	0.822	0.8

4. Conclusions

The aim of this study was to construct an SVR model for accelerating the design of ABO₃-type perovskites with optimal SSA, using a six-step approach. The analysis mainly focused on determining the influence of the key features on the SSA (1.05~56 m²g⁻¹) of ABO₃-type perovskites synthesized through the sol-gel technique. Results from LOOCV, external test, and the validation set demonstrate that the SVR-RBF model effectively predicts SSA of ABO₃-type perovskites with robust accuracy. Key features affecting SSA of ABO₃-type perovskites were identified using FSM and BSM incorporated in SVR-RBF. The perovskite candidates with higher SSA (67.884 m²g⁻¹) than the highest value (56 m²g⁻¹) of SSA observed using the existing dataset were identified by the SVR-RBF model, virtual screening, and pattern recognition. A webserver (http://1.14.49.110/online_predict/Perovskite (accessed on 7 June 2024)) based on the constructed model was developed to facilitate ease of use for researchers interested in ABO₃-type perovskites. This webserver allows rapid and convenient prediction of SSA of ABO₃-type perovskites and it is freely accessible. Furthermore, with the increase in relevant experimental data, we aim to collect more data and apply more algorithms to further enhance the predictive performance of the model in future endeavors. The approach outlined in this study is instrumental in accelerating material design and can be used for various applications in addition to ABO₃-type perovskite materials.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ma17123026/s1>, Table S1: Overview of ML applications in ABO₃-type perovskites; Table S2: The dataset including 102 ABO₃-type perovskite samples; Table S3: The meanings of the twenty-five candidate features; Figure S1: Pearson correlation matrix diagram of the twenty-five features; Figure S2 The mRMR scores of the twenty-five features; Formulas for 6 principal components. References [37,38,69–120] are cited in the Supplementary Materials.

Author Contributions: X.Z.: conceptualization, methodology, software, data curation, writing—original draft preparation. M.C.: visualization, investigation, supervision, validation, writing—reviewing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Sichuan Science and Technology Program of China (grant number 2022YFG0318), Panzhihua Instructional Science and Technology Program of China (grant number 2020ZD-G-11), and Panzhihua University Science and Technology Program of China grant number 2021PY009).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Materials, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors have no conflicts to declare.

Abbreviations

Abbreviation	Full Name
SSA	Specific Surface Area
ML	Machine Learning
SVR	Support Vector Regression
LOOCV	Leave-one-out Cross-validation
RBF	Radial Basis Function
<i>PM</i>	Synthetic Mode
<i>CT</i>	Calcination Temperature
<i>AH</i>	Calcination Time
<i>DT</i>	Drying Temperature
OCPMDM	Online Computational Platform of Material Data Mining
mRMR	Max Relevance Min Redundancy
R_A/R_B	Ratio of ionic radius
α	Unit cell lattice edge
t_f	Tolerance factor
R_A	Ionic Radius of A-site
R_B	Ionic Radius of B-site
R_O	Ionic Radius of O
DTR	Decision Tree Regression
GBR	Gradient Boosting Regression
PLS	Partial Least Squares
RVM	Relevance Vector Machine
LKF	Linear Kernel Function
PKF	Polynomial Kernel Function
RFR	Random Forest Regression
BPNN	Back Propagation Neural Network
RMSE	Root Mean Square Error
R	Pearson correlation coefficient
FSM	Forward Selection Method
BSM	Backward Selection Method
GA	Genetic Algorithm
Z_{IA}	Ionization Energy of A-site
Δfus_A	Enthalpy of fusion at the melting point of A-site
T_{mA}	Melting Point of A-site
T_{mB}	Melting Point of B-site
T_{bA}	Boiling Point of A-site
ρ_A	Density of A-site
EA_a	Electron Affinity of A-site
EA_b	Electron Affinity of B-site
P_{CA3}	Third principal component
P_{CA6}	Sixth principal component
R^2	The coefficient of determination

References

1. Ji, Q.Q.; Bi, L.; Zhang, J.T.; Cao, H.J.; Zhao, X.S. The role of oxygen vacancies of ABO₃ perovskite oxides in the oxygen reduction reaction. *Energy Environ. Sci.* **2020**, *13*, 1408–1428. [[CrossRef](#)]
2. Zhao, J.F.; Gao, J.C.; Li, W.M.; Qian, Y.T.; Shen, X.D.; Wang, X.; Shen, X.; Hu, Z.W.; Dong, C.; Huang, Q.Z.; et al. A combinatory ferroelectric compound bridging simple ABO₃ and A-site-ordered quadruple perovskite. *Nat. Commun.* **2021**, *12*, 747. [[CrossRef](#)]
3. Zhu, J.J.; Li, H.L.; Zhong, L.Y.; Xiao, P.; Xu, X.L.; Yang, X.G.; Zhao, Z.; Li, J.L. Perovskite oxides: Preparation, characterizations, and applications in heterogeneous catalysis. *ACS Catal.* **2014**, *4*, 2917–2940. [[CrossRef](#)]
4. Shang, C.Y.; Xiao, X.; Xu, Q. Coordination chemistry in modulating electronic structures of perovskite-type oxide nanocrystals for oxygen evolution catalysis. *Coord. Chem. Rev.* **2023**, *485*, 215109. [[CrossRef](#)]
5. Wang, Y.; Lv, Z.Y.; Zhou, L.; Chen, X.L.; Chen, J.R.; Zhou, Y.; Roy, V.A.L.; Han, S.T. Emerging perovskite materials for high density data storage and artificial synapses. *J. Mater. Chem. C* **2018**, *6*, 1600–1617. [[CrossRef](#)]
6. Wang, Y.N.; Wang, L.A.; Zhang, K.X.; Xu, J.Y.; Wu, Q.N.; Xie, Z.B.; An, W.; Liang, X.; Zou, X.X. Electrocatalytic water splitting over perovskite oxide catalysts. *Chin. J. Catal.* **2023**, *50*, 109–125. [[CrossRef](#)]
7. Kumar, A.; Kumar, A.; Krishnan, V. Perovskite oxide based materials for energy and environment-oriented photocatalysis. *ACS Catal.* **2020**, *10*, 10253–10315. [[CrossRef](#)]
8. Bhattar, S.; Abedin, M.A.; Kanitkar, S.; Spivey, J.J. A review on dry reforming of methane over perovskite derived catalysts. *Catal. Today* **2021**, *365*, 2–23. [[CrossRef](#)]
9. Sun, C.W.; Alonso, J.A.; Bian, J.J. Recent advances in perovskite-type oxides for energy conversion and storage applications. *Adv. Energy Mater.* **2021**, *11*, 2000459. [[CrossRef](#)]
10. Deeksha; Kour, P.; Ahmed, I.; Sunny; Sharma, S.K.; Yadav, K.; Mishra, Y.K. Transition metal-based perovskite oxides: Emerging electrocatalysts for oxygen evolution reaction. *ChemCatChem* **2023**, *15*, e202300040.
11. Zhang, S.J.; Jia, Z.R.; Cheng, B.; Zhao, Z.W.; Lu, F.; Wu, G.L. Recent progress of perovskite oxides and their hybrids for electromagnetic wave absorption: A mini-review. *Adv. Compos. Hybrid Mater.* **2022**, *5*, 2440–2460. [[CrossRef](#)]
12. Zhang, F.; Zhu, K. Additive engineering for efficient and stable perovskite solar cells. *Adv. Energy Mater.* **2020**, *10*, 1902579. [[CrossRef](#)]
13. Zuo, C.T.; Bolink, H.J.; Han, H.W.; Huang, J.S.; Cahen, D.; Ding, L.M. Advances in perovskite solar cells. *Adv. Sci.* **2016**, *3*, 1500324. [[CrossRef](#)] [[PubMed](#)]
14. Gharibzadeh, S.; Nejand, B.A.; Jakoby, M.; Abzieher, T.; Hauschild, D.; Moghadamzadeh, S.; Schwenzer, J.A.; Brenner, P.; Schmager, R.; Haghighirad, A.A.; et al. Record open-circuit voltage wide-bandgap perovskite solar cells utilizing 2D/3D perovskite heterostructure. *Adv. Energy Mater.* **2019**, *9*, 1803699. [[CrossRef](#)]
15. Chen, W.J.; Huang, Z.M.; Yao, H.T.; Liu, Y.; Zhang, Y.H.; Li, Z.J.; Zhou, H.M.; Xiao, P.; Chen, T.; Sun, H.D.; et al. Highly bright and stable single-crystal perovskite light-emitting diodes. *Nat. Photon.* **2023**, *17*, 401–407. [[CrossRef](#)]
16. Bai, S.; Wu, Z.W.; Wu, X.J.; Jin, Y.Z.; Zhao, N.; Chen, Z.H.; Mei, Q.Q.; Wang, X.Z.; Ye, Z.; Song, T.Y.; et al. High-performance planar heterojunction perovskite solar cells: Preserving long charge carrier diffusion lengths and interfacial engineering. *Nano Res.* **2014**, *7*, 1749–1758. [[CrossRef](#)]
17. Zhumekenov, A.A.; Saidaminov, M.I.; Haque, M.A.; Alarousu, E.; Sarmah, S.P.; Murali, B.; Dursun, I.; Miao, X.H.; Abdelhady, A.L.; Wu, T.; et al. Formamidinium lead halide perovskite crystals with unprecedented long carrier dynamics and diffusion length. *ACS Energy Lett.* **2016**, *1*, 32–37. [[CrossRef](#)]
18. Yettapu, G.R.; Talukdar, D.; Sarkar, S.; Swarnkar, A.; Nag, A.; Ghosh, P.; Mandal, P. Terahertz conductivity within colloidal CsPbBr₃ perovskite nanocrystals: Remarkably high carrier mobilities and large diffusion lengths. *Nano Lett.* **2016**, *16*, 4838–4848. [[CrossRef](#)] [[PubMed](#)]
19. Oga, H.; Saeki, A.; Ogomi, Y.; Hayase, S.; Seki, S. Improved understanding of the electronic and energetic landscapes of perovskite solar cells: High local charge carrier mobility, reduced recombination, and extremely shallow traps. *J. Am. Chem. Soc.* **2014**, *136*, 13818–13825. [[CrossRef](#)]
20. Jana, A.; Cho, S.; Patil, S.A.; Meena, A.; Jo, Y.; Sree, V.G.; Park, Y.; Kim, H.; Im, H.; Taylor, R.A. Perovskite: Scintillators, direct detectors, and X-ray imagers. *Mater. Today* **2022**, *55*, 110–136. [[CrossRef](#)]
21. Ma, F.; Zhao, Y.; Li, J.H.; Zhang, X.W.; Gu, H.S.; You, J.B. Nickel oxide for inverted structure perovskite solar cells. *J. Energy Chem.* **2021**, *52*, 393–411. [[CrossRef](#)]
22. Hossain, M.I.; Alharbi, F.H.; Tabet, N. Copper oxide as inorganic hole transport material for lead halide perovskite based solar cells. *Sol. Energy* **2015**, *120*, 370–380. [[CrossRef](#)]
23. Huang, L.; Huang, X.J.; Yan, J.; Liu, Y.H.; Jiang, H.; Zhang, H.G.; Tang, J.F.; Liu, Q. Research progresses on the application of perovskite in adsorption and photocatalytic removal of water pollutants. *J. Hazard. Mater.* **2023**, *442*, 130024. [[CrossRef](#)]
24. Li, C.L.; Han, C.; Zhang, Y.B.; Zang, Z.G.; Wang, M.; Tang, X.S.; Du, J.H. Enhanced photoresponse of self-powered perovskite photodetector based on ZnO nanoparticles decorated CsPbBr₃ films. *Sol. Energy Mater. Sol. Cells* **2017**, *172*, 341–346. [[CrossRef](#)]
25. Zhao, K.; He, F.; Huang, Z.; Wei, G.Q.; Zheng, A.Q.; Li, H.B.; Zhao, Z.L. Perovskite-type oxides LaFe_{1-x}Co_xO₃ for chemical looping steam methane reforming to syngas and hydrogen co-production. *Appl. Energy* **2016**, *168*, 193–203. [[CrossRef](#)]
26. Qin, F.; Chen, J.; Liu, J.W.; Liu, L.; Tang, C.J.; Tang, B.; Li, G.F.; Zeng, L.C.; Li, H.L.; Yi, Z. Design of high efficiency perovskite solar cells based on inorganic and organic undoped double hole layer. *Sol. Energy* **2023**, *262*, 111796. [[CrossRef](#)]

27. Qin, W.B.; Yuan, Z.Y.; Gao, H.L.; Zhang, R.Z.; Meng, F.L. Perovskite-structured LaCoO₃ modified ZnO gas sensor and investigation on its gas sensing mechanism by first principle. *Sens. Actuators B Chem.* **2021**, *341*, 130015. [[CrossRef](#)]
28. Hossain, M.K.; Toki, G.F.I.; Kuddus, A.; Rubel, M.H.K.; Hossain, M.M.; Bencherif, H.; Rahman, M.F.; Islam, M.R.; Mushtaq, M. An extensive study on multiple ETL and HTL layers to design and simulation of high-performance lead-free CsSnCl₃-based perovskite solar cells. *Sci. Rep.* **2023**, *13*, 2521. [[CrossRef](#)]
29. Luo, W.; Phung, D.; Tran, T.; Gupta, S.; Rana, S.; Karmakar, C.; Shilton, A.; Yearwood, J.; Dimitrova, N.; Ho, T.B.; et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: A multidisciplinary view. *J. Med. Internet Res.* **2016**, *18*, e323. [[CrossRef](#)] [[PubMed](#)]
30. Cuoco, E.; Powell, J.; Cavaglià, M.; Ackley, K.; Bejger, M.; Chatterjee, C.; Coughlin, M.; Coughlin, S.; Easter, P.; Essick, R.; et al. Enhancing gravitational-wave science with machine learning. *Mach. Learn. Sci. Technol.* **2021**, *2*, 011002. [[CrossRef](#)]
31. Hamdia, K.M.; Ghasemi, H.; Bazi, Y.; AlHichri, H.; Alajlan, N.; Rabczuk, T. A novel deep learning based method for the computational material design of flexoelectric nanostructures with topology optimization. *Finite Elem. Anal. Des.* **2019**, *165*, 21–30. [[CrossRef](#)]
32. Fu, Y.Z.; Downey, A.R.J.; Yuan, L.; Zhang, T.Y.; Pratt, A.; Balogun, Y. Machine learning algorithms for defect detection in metal laser-based additive manufacturing: A review. *J. Manuf. Process.* **2022**, *75*, 693–710. [[CrossRef](#)]
33. Zhang, L.; He, M.; Shao, S.F. Machine learning for halide perovskite materials. *Nano Energy* **2020**, *78*, 105380. [[CrossRef](#)]
34. Wang, Z.M.; Yang, M.; Xie, X.X.; Yu, C.Y.; Jiang, Q.L.; Huang, M.N.; Algadi, H.; Guo, Z.H.; Zhang, H. Applications of machine learning in perovskite materials. *Adv. Compos. Hybrid Mater.* **2022**, *5*, 2700–2720. [[CrossRef](#)]
35. Ali, A.; Park, H.; Mall, R.; Aissa, B.; Sanvito, S.; Bensmail, H.; Belaidi, A.; El-Mellouhi, F. Machine learning accelerated recovery of the cubic structure in mixed-cation perovskite thin films. *Chem. Mater.* **2020**, *32*, 2998–3006. [[CrossRef](#)]
36. Weng, B.C.; Song, Z.L.; Zhu, R.L.; Yan, Q.Y.; Sun, Q.D.; Grice, C.G.; Yan, Y.F.; Yin, W.J. Simple descriptor derived from symbolic regression accelerating the discovery of new perovskite catalysts. *Nat. Commun.* **2020**, *11*, 3513. [[CrossRef](#)] [[PubMed](#)]
37. Tao, Q.L.; Chang, D.P.; Lu, T.; Li, L.; Chen, H.M.; Yang, X.; Liu, X.J.; Li, M.J.; Lu, W.C. Multiobjective stepwise design strategy-assisted design of high-performance perovskite oxide photocatalysts. *J. Phys. Chem. C* **2021**, *125*, 21141–21150. [[CrossRef](#)]
38. Shi, L.; Chang, D.P.; Ji, X.B.; Lu, W.C. Using data mining to search for perovskite materials with higher specific surface area. *J. Chem. Inf. Model.* **2018**, *58*, 2420–2427. [[CrossRef](#)]
39. Yang, J.Q.; Mannodi-Kanakkithodi, A. High-throughput computations and machine learning for halide perovskite discovery. *MRS Bull.* **2022**, *47*, 940–948. [[CrossRef](#)]
40. Wei, K.X.; Faraj, Y.; Yao, G.; Xie, R.Z.; Lai, B. Strategies for improving perovskite photocatalysts reactivity for organic pollutants degradation: A review on recent progress. *Chem. Eng. J.* **2021**, *414*, 128783. [[CrossRef](#)]
41. Bajorowicz, B.; Nadolna, J.; Lisowski, W.; Klimczuk, T.; Zaleska-Medynska, A. The effects of bifunctional linker and reflux time on the surface properties and photocatalytic activity of CdTe quantum dots decorated KTaO₃ composite photocatalysts. *Appl. Catal. B Environ.* **2017**, *203*, 452–464. [[CrossRef](#)]
42. Ma, Y.J.; Su, P.; Ge, Y.Z.; Wang, F.W.; Xue, R.X.; Wang, Z.J.; Li, Y.S. A novel LaAlO₃ perovskite with large surface area supported Ni-based catalyst for methane dry reforming. *Catal. Lett.* **2022**, *152*, 2993–3003. [[CrossRef](#)]
43. Sohrabian, M.; Mahdikhah, V.; Alimohammadi, E.; Sheibani, S. Improved photocatalytic performance of SrTiO₃ through a Z-scheme polymeric-perovskite heterojunction with g-C₃N₄ and plasmonic resonance of Ag mediator. *Appl. Surf. Sci.* **2023**, *618*, 156682. [[CrossRef](#)]
44. Zulfiqar, W.; Javed, F.; Abbas, G.; Larsson, J.A.; Alay-e-Abbas, S.M. Stabilizing the dopability of chalcogens in BaZrO₃ through TiZr co-doping and its impact on the opto-electronic and photocatalytic properties: A meta-GGA level DFT study. *Int. J. Hydrogen Energy* **2024**, *58*, 409–415. [[CrossRef](#)]
45. Hagendorff, T. Linking human and machine behavior: A new approach to evaluate training data quality for beneficial machine learning. *Minds Mach.* **2021**, *31*, 563–593. [[CrossRef](#)] [[PubMed](#)]
46. Akhter, R.; Sofi, S.A. Precision agriculture using IoT data analytics and machine learning. *J. King Saud Univ. Comput. Inf. Sci.* **2022**, *34*, 5602–5618. [[CrossRef](#)]
47. Qin, W.B.; Yuan, Z.Y.; Shen, Y.B.; Zhang, R.Z.; Meng, F.L. Phosphorus-doped porous perovskite LaFe_{1-x}P_xO_{3-δ} nanosheets with rich surface oxygen vacancies for ppb level acetone sensing at low temperature. *Chem. Eng. J.* **2022**, *431*, 134280. [[CrossRef](#)]
48. Chang, D.P.; Lu, W.C.; Wang, G. Designing bulk metallic glasses materials with higher reduced glass transition temperature via machine learning. *Chemom. Intell. Lab. Syst.* **2022**, *228*, 104621. [[CrossRef](#)]
49. Chang, D.P.; Xu, P.C.; Ji, X.B.; Li, M.J.; Lu, W.C. Application of Online Computational Platform of Materials Data Mining (OCPMDM) in Search for ABO₃ Perovskites with Multi-Properties. *Sci. Adv. Mater.* **2023**, *15*, 1014–1025. [[CrossRef](#)]
50. Xu, P.C.; Ji, X.B.; Li, M.J.; Lu, W.C. Small data machine learning in materials science. *NPJ Comput. Mater.* **2023**, *9*, 42. [[CrossRef](#)]
51. Chang, D.P.; Xu, P.C.; Li, M.J.; Lu, W.C. OCPMDM 2.0: An intelligent solution for materials data mining. *Chemom. Intell. Lab. Syst.* **2023**, *243*, 105022.
52. Shen, Y.S.; Wang, J.Y.; Ji, X.B.; Lu, W.C. Machine Learning-Assisted Discovery of 2D Perovskites with Tailored Bandgap for Solar Cells. *Adv. Theory Simul.* **2023**, *6*, 2200922. [[CrossRef](#)]
53. Liu, X.J.; Xu, P.C.; Zhao, J.J.; Lu, W.C.; Li, M.J.; Wang, G. Material machine learning for alloys: Applications, challenges and perspectives. *J. Alloys Compd.* **2022**, *921*, 165984. [[CrossRef](#)]

54. Roy, A.; Chakraborty, S. Support vector machine in structural reliability analysis: A review. *Reliab. Eng. Syst. Saf.* **2023**, *233*, 109126. [[CrossRef](#)]
55. Huang, Y.; Zhao, L. Review on landslide susceptibility mapping using support vector machines. *Catena* **2018**, *165*, 520–529. [[CrossRef](#)]
56. Handelman, G.S.; Kok, H.K.; Chandra, R.V.; Razavi, A.H.; Huang, S.W.; Brooks, M.; Lee, M.J.; Asadi, H. Peering into the black box of artificial intelligence: Evaluation metrics of machine learning methods. *Am. J. Roentgenol.* **2019**, *212*, 38–43. [[CrossRef](#)] [[PubMed](#)]
57. Gui, G.; Liu, F.; Sun, J.L.; Yang, J.; Zhou, Z.Q.; Zhao, D.X. Flight delay prediction based on aviation big data and machine learning. *IEEE Trans. Veh. Technol.* **2020**, *69*, 140–150. [[CrossRef](#)]
58. Li, J.D.; Cheng, K.W.; Wang, S.H.; Morstatter, F.; Trevino, R.P.; Tang, J.L.; Liu, H. Feature selection: A data perspective. *ACM Comput. Surv.* **2018**, *50*, 1–45. [[CrossRef](#)]
59. He, Z.M.; Li, L.Z.; Huang, Z.M.; Situ, H.Z. Quantum-enhanced feature selection with forward selection and backward elimination. *Quantum Inf. Process.* **2018**, *17*, 154. [[CrossRef](#)]
60. Aregbesola, S.O.; Won, J.; Kim, S.; Byun, Y.H. Sequential backward feature selection for optimizing permanent strain model of unbound aggregates. *Case Stud. Constr. Mater.* **2023**, *19*, e02554. [[CrossRef](#)]
61. Oreski, S.; Oreski, G. Genetic algorithm-based heuristic for feature selection in credit risk assessment. *Expert Syst. Appl.* **2014**, *41*, 2052–2064. [[CrossRef](#)]
62. Zhang, Y.; Gong, D.W.; Gao, X.Z.; Tian, T.; Sun, X.Y. Binary differential evolution with self-learning for multi-objective feature selection. *Inf. Sci.* **2020**, *507*, 67–85. [[CrossRef](#)]
63. Ait-Sahalia, Y.; Xiu, D.C. Principal Component Analysis of High-Frequency Data. *J. Am. Stat. Assoc.* **2019**, *114*, 287–303. [[CrossRef](#)]
64. Navarro, C.L.A.; Damen, J.A.A.; Takada, T.; Nijman, S.W.J.; Dhiman, P.; Ma, J.; Collins, G.S.; Bajpai, R.; Riley, R.D.; Moons, K.G.M.; et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: Systematic review. *BMJ Br. Med. J.* **2021**, *375*, n2281. [[CrossRef](#)]
65. Lindauer, M.; Eggensperger, K.; Feurer, M.; Biedenkapp, A.; Deng, D.; Benjamins, C.; Ruhkopf, T.; Sass, R.; Hutter, F. SMAC3: A versatile bayesian optimization package for hyperparameter optimization. *J. Mach. Learn. Res.* **2022**, *23*, 1–9.
66. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [[CrossRef](#)]
67. Mahmoudi, N.; Duman, E. Detecting credit card fraud by Modified Fisher Discriminant Analysis. *Expert Syst. Appl.* **2015**, *42*, 2510–2516. [[CrossRef](#)]
68. Zheng, Z.; Petrone, R.; Péra, M.C.; Hissel, D.; Becherif, M.; Pianese, C.; Steiner, N.Y.; Sorrentino, M. A review on non-model based diagnosis methodologies for PEM fuel cell stacks and systems. *Int. J. Hydrogen Energy* **2013**, *38*, 8914–8926. [[CrossRef](#)]
69. Li, W.; Jacobs, R.; Morgan, D. Predicting the thermodynamic stability of perovskite oxides using machine learning models. *Comp. Mater. Sci.* **2018**, *150*, 454–463. [[CrossRef](#)]
70. Liu, H.Y.; Cheng, J.C.; Dong, H.Z.; Feng, J.G.; Pang, B.L.; Tian, Z.Y.; Ma, S.; Xia, F.J.; Zhang, C.K.; Dong, L.F. Screening stable and metastable ABO₃ perovskites using machine learning and the materials project. *Comp. Mater. Sci.* **2020**, *177*, 109614. [[CrossRef](#)]
71. Balachandran, P.V.; Emery, A.A.; Gubernatis, J.E.; Lookman, T.; Wolverton, C.; Zunger, A. Predictions of new ABO₃ perovskite compounds by combining machine learning and density functional theory. *Phys. Rev. Mater.* **2018**, *2*, 043802. [[CrossRef](#)]
72. Li, C.J.; Hao, H.; Xu, B.; Zhao, G.H.; Chen, L.H.; Zhang, S.J.; Liu, H.X. A progressive learning method for predicting the band gap of ABO₃ perovskites using an instrumental variable. *J. Mater. Chem. C* **2020**, *8*, 3127–3136. [[CrossRef](#)]
73. Zhai, X.Y.; Chen, M.T.; Lu, W.C. Accelerated search for perovskite materials with higher Curie temperature based on the machine learning methods. *Comp. Mater. Sci.* **2018**, *151*, 41–48. [[CrossRef](#)]
74. Yang, Z.X.; Gao, Z.R.; Sun, X.F.; Cai, H.L.; Zhang, F.M.; Wu, X.S. High critical transition temperature of lead-based perovskite ferroelectric crystals: A machine learning study. *ACTA Phys. Sin-Ch Ed.* **2019**, *68*, 210502. [[CrossRef](#)]
75. Xiao, L.H.; Zhang, Q.; Xu, X.; Ji, X.B.; Lu, W.C. Support vector regression assisted predictions the néel temperature of perovskites manganites. *Comput. Appl. Chem.* **2018**, *35*, 349–357.
76. Zhang, Y.; Xu, X.J. Machine learning the magnetocaloric effect in manganites from lattice parameters. *Appl. Phys. A Mater.* **2020**, *126*, 341. [[CrossRef](#)]
77. Xu, L.; Wencong, L.; Chunrong, P.; Qiang, S.; Jin, G. Two semi-empirical approaches for the prediction of oxide ionic conductivities in ABO₃ perovskites. *Comp. Mater. Sci.* **2009**, *46*, 860–868. [[CrossRef](#)]
78. Priyanga, G.S.; Mattur, M.N.; Nagappan, N.; Rath, S.; Thomas, T. Prediction of nature of band gap of perovskite oxides (ABO₃) using a machine learning approach. *J. Mater.* **2022**, *8*, 937–948.
79. Chen, L.P.; Wang, X.C.; Xia, W.J.; Liu, C.H. PSO-SVR predicting for the Ehull of ABO₃-type compounds to screen the thermodynamic stable perovskite candidates based on multi-scale descriptors. *Comp. Mater. Sci.* **2022**, *211*, 111435. [[CrossRef](#)]
80. Priyadarshini, R.; Joardar, H.; Bisoy, S.K.; Badapanda, T. Crystal structural prediction of perovskite materials using machine learning: A comparative study. *Solid. State Commun.* **2023**, *361*, 115062. [[CrossRef](#)]
81. Naseri, M.; Gusarov, S.; Salahub, D.R. Quantum Machine Learning in Materials Prediction: A Case Study on ABO₃ Perovskite Structures. *J. Phys. Chem. Lett.* **2023**, *14*, 6940–6947. [[CrossRef](#)]
82. Chen, L.P.; Xia, W.J.; Yao, T.Z. Identifying descriptors for perovskite structure of composite oxides and inferring formability via low-dimensional described features. *Comp. Mater. Sci.* **2023**, *226*, 112216. [[CrossRef](#)]

83. Fan, X.Y. Prediction of formation energy using two-stage machine learning based on clustering. *Mater. Tehmol.* **2021**, *55*, 263–268. [[CrossRef](#)]
84. Sharma, V.; Kumar, P.; Dev, P.; Pilania, G. Machine learning substitutional defect formation energies in ABO_3 perovskites. *J. Appl. Phys.* **2020**, *128*. [[CrossRef](#)]
85. McGuinness, K.P.; Oliynyk, A.O.; Lee, S.; Molero-Sanchez, B.; Addo, P.K. Machine-learning prediction of thermal expansion coefficient for perovskite oxides with experimental validation. *Phys. Chem. Chem. Phys.* **2023**, *25*, 32123–32131. [[CrossRef](#)]
86. Zhang, Y.; Xu, X.J. Modeling oxygen ionic conductivities of ABO_3 Perovskites through machine learning. *Chem. Phys.* **2022**, *558*, 111511. [[CrossRef](#)]
87. Zhao, J.; Wang, X.Y. Screening perovskites from ABO_3 combinations generated by constraint satisfaction techniques using machine learning. *ACS Omega* **2022**, *7*, 10483–10491. [[CrossRef](#)]
88. Lourenço, M.P.; Tchagang, A.; Shankar, K.; Thangadurai, V.; Salahub, D.R. Active learning for optimum experimental design—insight into perovskite oxides. *Can. J. Chem.* **2023**, *101*, 734–744. [[CrossRef](#)]
89. Bhattacharya, S.; Roy, A. Linking stability with molecular geometries of perovskites and lanthanide richness using machine learning methods. *Comp. Mater. Sci.* **2024**, *231*, 112581. [[CrossRef](#)]
90. Li, W.; Wang, Z.G.; Xiao, X.; Zhang, Z.Q.; Janotti, A.; Rajasekaran, S.; Medasani, B. Predicting band gaps and band-edge positions of oxide perovskites using density functional theory and machine learning. *Phys. Rev. B* **2022**, *106*, 155156. [[CrossRef](#)]
91. Perween, S.; Ranjan, A. Improved visible-light photocatalytic activity in $ZnTiO_3$ nanopowder prepared by sol-electrospinning. *Sol. Energ. Mat. Sol. C* **2017**, *163*, 148–156. [[CrossRef](#)]
92. Orak, C.; Atalay, S.; Ersöz, G. Photocatalytic and photo-Fenton-like degradation of methylparaben on monolith-supported perovskite-type catalysts. *Sep. Sci. Technol.* **2017**, *52*, 1310–1320. [[CrossRef](#)]
93. Sun, H.; Yang, H.; Cui, S.; Nie, K.; Wu, J. Simultaneous Mg-modification inside and outside of $LaCoO_3$ lattice and their photocatalytic properties. *Chin. J. Inorg. Chem.* **2016**, *32*, 1704–1712.
94. Abdulkadir, I.; Jonnalagadda, S.B.; Martincigh, B.S. Synthesis and effect of annealing temperature on the structural, magnetic and photocatalytic properties of $(La_{0.5}Bi_{0.2}Ba_{0.2}Mn_{0.1})FeO_{(3-\delta)}$. *Mater. Chem. Phys.* **2016**, *178*, 196–203. [[CrossRef](#)]
95. Josephine, B.A.; Manikandan, A.; Teresita, V.M.; Antony, S.A. Fundamental study of $LaMg_xCr_{1-x}O_{3-\delta}$ perovskites nano-photocatalysts: Sol-gel synthesis, characterization and humidity sensing. *Korean, J. Chem. Eng.* **2016**, *33*, 1590–1598. [[CrossRef](#)]
96. Tijare, S.N.; Bakardjieva, S.; Subrt, J.; Joshi, M.V.; Rayalu, S.S.; Hishita, S.; Labhsetwar, N. Synthesis and visible light photocatalytic activity of nanocrystalline $PrFeO_3$ perovskite for hydrogen generation in ethanol-water system. *J. Chem. Sci.* **2014**, *126*, 517–525. [[CrossRef](#)]
97. Tavakkoli, H.; Yazdanbakhsh, M. Fabrication of two perovskite-type oxide nanoparticles as the new adsorbents in efficient removal of a pesticide from aqueous solutions: Kinetic, thermodynamic, and adsorption studies. *Micropor. Mesopor. Mater.* **2013**, *176*, 86–94. [[CrossRef](#)]
98. Li, H.; Cui, Y.; Wu, X.; Hong, W.; Hua, L. Effect of La contents on the structure and photocatalytic activity of $La-SrTiO_3$ catalysts. *Chin. J. Inorg. Chem.* **2012**, *28*, 2597–2604.
99. Li, Y.; Yao, S.; Wen, W.; Xue, L.; Yan, Y. Sol-gel combustion synthesis and visible-light-driven photocatalytic property of perovskite $LaNiO_3$. *J. Alloy Compd.* **2010**, *491*, 560–564. [[CrossRef](#)]
100. Li, S.; Jing, L.; Fu, W.; Yang, L.; Xin, B.; Fu, H. Photoinduced charge property of nanosized perovskite-type $LaFeO_3$ and its relationships with photocatalytic activity under visible irradiation. *Mater. Res. Bull.* **2007**, *42*, 203–212. [[CrossRef](#)]
101. Puangpetch, T.; Sreethawong, T.; Chavadej, S. Hydrogen production over metal-loaded mesoporous-assembled $SrTiO_3$ nanocrystal photocatalysts: Effects of metal type and loading. *Int. J. Hydrogen Energy* **2010**, *35*, 6531–6540. [[CrossRef](#)]
102. Puangpetch, T.; Sreethawong, T.; Yoshikawa, S.; Chavadej, S. Hydrogen production from photocatalytic water splitting over mesoporous-assembled $SrTiO_3$ nanocrystal-based photocatalysts. *J. Mol. Catal. A-Chem.* **2009**, *312*, 97–106. [[CrossRef](#)]
103. Touahra, F.; Rabahi, A.; Chebout, R.; Boudjemaa, A.; Lerari, D.; Sehaïlia, M.; Halliche, D.; Bachari, K. Enhanced catalytic behaviour of surface dispersed nickel on $LaCuO_3$ perovskite in the production of syngas: An expedient approach to carbon resistance during CO_2 reforming of methane. *Int. J. Hydrogen Energy* **2016**, *41*, 2477–2486. [[CrossRef](#)]
104. Sutthiumporn, K.; Maneerung, T.; Kathiraser, Y.; Kawi, S. CO_2 dry-reforming of methane over $La_{0.8}Sr_{0.2}Ni_{0.8}M_{0.2}O_3$ perovskite ($M = Bi, Co, Cr, Cu, Fe$): Roles of lattice oxygen on C–H activation and carbon suppression. *Int. J. Hydrogen Energy* **2012**, *37*, 11195–11207. [[CrossRef](#)]
105. Luo, Y.; Wang, X.; Qian, Q.; Chen, Q. Studies on B sites in Fe-doped $LaNiO_3$ perovskite for SCR of NO_x with H_2 . *Int. J. Hydrogen Energy* **2014**, *39*, 15836–15843. [[CrossRef](#)]
106. Parida, K.M.; Reddy, K.H.; Martha, S.; Das, D.P.; Biswal, N. Fabrication of nanocrystalline $LaFeO_3$: An efficient sol-gel auto-combustion assisted visible light responsive photocatalyst for water decomposition. *Int. J. Hydrogen Energy* **2010**, *35*, 12161–12168. [[CrossRef](#)]
107. Dhanasekaran, P.; Gupta, N.M. Factors affecting the production of H_2 by water splitting over a novel visible-light-driven photocatalyst $GaFeO_3$. *Int. J. Hydrogen Energy* **2012**, *37*, 4897–4907. [[CrossRef](#)]
108. Husin, H.; Chen, H.M.; Su, W.N.; Pan, C.J.; Chuang, W.T.; Sheu, H.-S.; Hwang, B.-J. Green fabrication of La-doped $NaTaO_3$ via H_2O_2 assisted sol-gel route for photocatalytic hydrogen production. *Appl. Catal. B Environ. Energy* **2011**, *102*, 343–351. [[CrossRef](#)]
109. Bui, D.-N.; Mu, J.; Wang, L.; Kang, S.-Z.; Li, X. Preparation of Cu-loaded $SrTiO_3$ nanoparticles and their photocatalytic activity for hydrogen evolution from methanol aqueous solution. *Appl. Surf. Sci.* **2013**, *274*, 328–333. [[CrossRef](#)]

110. Zhang, H.; Chen, G.; He, X.; Xu, J. Electronic structure and photocatalytic properties of Ag–La codoped CaTiO₃. *J. Alloy Compd.* **2012**, *516*, 91–95. [[CrossRef](#)]
111. Tijare, S.N.; Joshi, M.V.; Padole, P.S.; Mangrulkar, P.A.; Rayalu, S.S.; Labhsetwar, N.K. Photocatalytic hydrogen generation through water splitting on nano-crystalline LaFeO₃ perovskite. *Int. J. Hydrogen Energy* **2012**, *37*, 10451–10456. [[CrossRef](#)]
112. Teresita, V.M.; Manikandan, A.; Josephine, B.A.; Sujatha, S.; Antony, S.A. Electromagnetic properties and humidity-sensing studies of magnetically recoverable LaMg_xFe_{1-x}O_{3-δ} perovskites nano-photocatalysts by Sol-Gel route. *J. Supercond. Nov. Magn.* **2016**, *29*, 1691–1701. [[CrossRef](#)]
113. Sydorchuk, V.; Lutsyuk, I.; Shved, V.; Hreb, V.; Kondyr, A.; Zakutevskyy, O.; Vasylechko, L. PrCo_{1-x}FexO₃ perovskite powders for possible photocatalytic applications. *Res. Chem. Intermediat* **2020**, *46*, 1909–1930. [[CrossRef](#)]
114. Ibarra-Rodriguez, L.I.; Huerta-Flores, A.M.; Mora-Hernandez, J.M.; Torres-Martínez, L.M. Photocatalytic evolution of H₂ over visible-light active LaMO₃(M: Co, Mn, Fe) perovskite materials: Roles of oxygenated species in catalytic performance. *J. Phys. Chem. Solids* **2020**, *136*, 109189. [[CrossRef](#)]
115. Mota, N.; Álvarez-Galván, M.C.; Al-Zahrani, S.M.; Navarro, R.M.; Fierro, J.L.G. Diesel fuel reforming over catalysts derived from LaCo_{1-x}Ru_xO₃ perovskites with high Ru loading. *Int. J. Hydrogen Energy* **2012**, *37*, 7056–7066. [[CrossRef](#)]
116. Agüero, F.N.; Morales, M.R.; Larrégola, S.; Izurieta, E.M.; Lopez, E.; Cadús, L.E. La_{1-x}Ca_xAl_{1-y}Ni_yO₃ perovskites used as precursors of nickel based catalysts for ethanol steam reforming. *Int. J. Hydrogen Energy* **2015**, *40*, 15510–15520. [[CrossRef](#)]
117. Hu, R.; Li, C.; Wang, X.; Sun, Y.; Jia, H.; Su, H.; Zhang, Y. Photocatalytic activities of LaFeO₃ and La₂FeTiO₆ in p-chlorophenol degradation under visible light. *Catal. Commun.* **2012**, *29*, 35–39. [[CrossRef](#)]
118. Boumaza, S.; Boudjellal, L.; Brahim, R.; Belhadi, A.; Trari, M. Synthesis by citrates sol-gel method and characterization of the perovskite LaFeO₃: Application to oxygen photo-production. *J. Sol-Gel Sci. Technol.* **2020**, *94*, 486–492. [[CrossRef](#)]
119. Wang, J.; Li, X.D.; Yu, Z.Q.; Zhang, S. Enhancing infrared emissivity of GdCoO₃ with Ca doping: Potential for advanced thermal control materials. *Ceram. Int.* **2024**, *50*, 9630–9639. [[CrossRef](#)]
120. Guo, J.X.; Jing, Y.; Shen, T.; Luo, H.D.; Liang, J.; Yuan, S.D. Effect of doped strontium on catalytic properties of La_{1-x}Sr_xMnO₃ for rhodamine B degradation. *J. Rare Earth* **2021**, *39*, 1362–1369. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.