Machine learning-enhanced design of lead-free halide perovskite materials using density functional theory

Upendra Kumar [a,1], Hyeon Woo Kim [a,b,1], Gyanendra Kumar Maurya [c], Bincy Babu Raj [a,d], Sobhit Singh [e,f], Ajay Kumar Kushwaha [g], Sung Beom Cho [h,i,*], Hyunseok Ko [a,*]

[a] *Division of Carbon Neutrality and Digitalization, Korea Institute of Ceramic Engineering and Technology (KICET), Jinju 52851, South Korea*
[b] *Department of Materials Science and Engineering, Hanyang University, Seoul 04763, South Korea*
[c] *Department of Physics, GLA University, Mathura 281406, India*
[d] *School of Material Science and Engineering, Pusan National University, 2 Busandaehak-ro 63beon-gil, Geumjeong-gu, Busan 46241, South Korea*
[e] *Department of Mechanical Engineering, University of Rochester, Rochester, NY 14627, USA*
[f] *Materials Science Program, University of Rochester, Rochester, NY 14627, USA*
[g] *Department of Metallurgy Engineering and Materials Science, Indian Institute of Technology Indore, Khandwa Road, Simrol, Indore 453552, India*
[h] *Department of Materials Science and Engineering, Ajou University, Suwon 16499, South Korea*
[i] *Department of Energy Systems Research, Ajou University, Suwon 16499, South Korea*

ARTICLE INFO

ABSTRACT

The investigation of emerging non-toxic perovskite materials has been undertaken to advance the fabrication of environmentally sustainable lead-free perovskite solar cells. This study introduces a machine learning methodology aimed at predicting innovative halide perovskite materials that hold promise for use in photovoltaic applications. The seven newly predicted materials are as follows: $CsMnCl_4$, $Rb_3Mn_2Cl_9$, Rb_4MnCl_6, Rb_3MnCl_5, $RbMn_2Cl_7$, $RbMn_4Cl_9$, and $CsIn_2Cl_7$. The predicted compounds are first screened using a machine learning approach, and their validity is subsequently verified through density functional theory calculations. $CsMnCl_4$ is notable among them, displaying a bandgap of 1.37 eV, falling within the Shockley-Queisser limit, making it suitable for photovoltaic applications. Through the integration of machine learning and density functional theory, this study presents a methodology that is more effective and thorough for the discovery and design of materials.

1. Introduction

Halide perovskites have emerged as promising candidates for revolutionizing the landscape of solar cell technology [1]. Characterized by their unique crystalline structure and exceptional optoelectronic properties, halide perovskites offer unparalleled potential for efficient and cost-effective solar energy conversion. Their tunable bandgap [2], high absorption coefficients [3], and long carrier diffusion lengths [4] make them highly attractive for photovoltaic applications. Moreover, their facile synthesis methods [5] and compatibility with flexible substrates [6] open avenues for scalable and versatile solar cell designs. Beyond their application in solar cells, halide perovskite materials hold immense promise for a wide range of electronic devices [7]. However, concerns over the toxicity of lead-based halide perovskites have prompted significant research efforts towards developing lead-free alternatives.

Lead-free perovskite halide materials [8] represent a promising avenue in the quest for sustainable and environmentally friendly solar energy solutions. These materials offer potential for high-performance solar cells [9] while mitigating environmental and health risks. Research efforts have focused on exploring a diverse range of lead-free perovskite compositions, including tin, bismuth, and other metal halides [10], which exhibit encouraging optoelectronic properties and demonstrate potential for efficient photovoltaic applications [11]. By addressing the need for lead-free halide designs, researchers aim to unlock the full potential of halide perovskite materials while ensuring environmental sustainability and human health safety.

Since only a few elements around lead in the periodic table have been considered, there are still many unexplored compositional regions in lead-free perovskites. This is due to the difficulty of experimental trial and error. Therefore, there is a need for techniques that can accelerate

material design, such as machine learning (ML). Research efforts have focused on synthesizing lead-free perovskite halides with enhanced stability, efficiency, and device performance [12]. Despite challenges in achieving commercial viability and scalability, ongoing advancements in material engineering and device fabrication techniques are paving the way for the integration of lead-free perovskite halides into next-generation electronic devices [13].

Recent efforts have been made to identify lead-free halide perovskite materials through data mining from the Materials Project [14]. The Debye temperature and thermal stability also have implications for thermal management in electronic devices [15]. By selecting materials with appropriate Debye temperatures, engineers can design electronic components that operate efficiently within specified temperature ranges without experiencing degradation or failure. This aspect is particularly crucial in the development of high-performance computing systems and advanced electronics. Hence, we compute the Debye temperature for the dataset comprising halide perovskites. Within this investigation, we have uncovered overlooked halide perovskite and predicted novel compounds suitable for photovoltaic purposes. ML techniques [16] are utilized for the prediction of vast number of halide perovskite materials. To confirm this prediction of the forecasted halide perovskite materials by ML, density functional theory (DFT) is utilized. This integrated method of combining ML with DFT offers a greater opportunity for streamlined and thorough method for the discovery and design of materials, aiding in the recognition of halide perovskite materials possessing sought-after electronic and optical properties suitable for photovoltaic applications.

2. Methods

ML Technique
The generalized Crystal Graph Convolutional Neural Networks (CGCNN) framework, developed by T. Xie and J. C. Grossman [16], is tailored to effectively represent periodic crystal systems. They [16] utilized eight distinct properties in their research: **(i)** crystallographic information, **(ii)** formation energy, **(iii)** absolute energy, **(iv)** bandgap, **(v)** Fermi energy, **(vi)** bulk moduli, **(vii)** shear moduli, and **(viii)** Poisson ratio. In their CGCNN model, crystallographic information was used as a feature, while the remaining seven properties served as targets. In the CGCNN model, structures are represented as graphs where the nodes correspond to atoms and the edges represent the bonds between them. The features used for these nodes are derived from elemental properties, which include atomic number, electronegativity, ionization energy, atomic radius, and other fundamental attributes of the elements. These features provide critical information that captures the chemical and physical characteristics of each atom, enabling the model to learn complex relationships and interactions within the crystal structure for accurate property prediction. The CGCNN model is trained via Crystallographic Information Files (CIF). A CIF is a standard text file format used to describe the crystallographic structure of a material. It contains detailed information about the atomic coordinates, symmetry operations, unit cell parameters, and other essential data needed to define the crystal structure. CIFs are widely used in materials science and crystallography to facilitate data exchange and enable accurate structure determination and analysis. Remarkably, CGCNN's interpretability is exhibited through the extraction of energy contributions specific to sites within perovskite structures, offering insights into how local chemical environments influence materials properties.

Dataset and Features
We have effectively obtained an extensive dataset from the Materials Project [14]. Hence, we conducted extensive data mining for halogen ternary compounds (perovskite halide), specifically involving Br, Cl, and I, leveraging the vast Materials Project database [14]. In the collated dataframe includes crucial physical quantities such as **(i)** material identification, **(ii)** pretty-formula representation, **(iii)** bandgap values, **(iv)** convex hull information, **(v)** space-group symbols, **(vi)** Inorganic Crystal
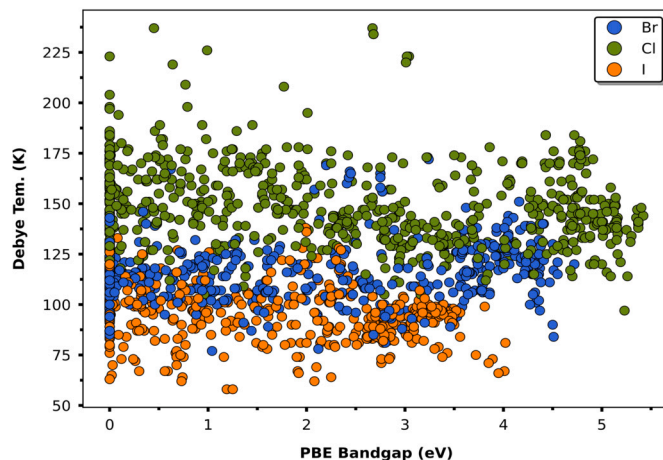

Fig. 1. The plot illustrates the correlation between Debye temperature and bandgap for ternary perovskite halides based on Br, Cl, and I. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

Structure Database (ICSD) identifiers, **(vii)** structural details, **(viii)** elasticity characteristics, **(xi)** Crystallographic Information File (CIF) data, and **(x)** Debye temperature values (Θ_D). The Θ_D and bandgap dataset encompasses a total of 456 and 2608 rows, respectively. Our objective was to swiftly predict key material properties i.e. Θ_D and bandgap for of lead-free halide perovskites. Both datasets consist of 10 columns, rendering it a valuable asset for conducting varied research and analyses in the realm of materials science. The CGCNN model's features stem from the crystal graph, depicted through crystallographic information files, offering a thorough portrayal of the structural intricacies of the crystal system. Approximately 70% of the data is designated for training. Another 15% is reserved for testing, and an equal 15% is allocated for validation. The CGCNN model undergoes a total of 30 epochs.

For screening materials, we used three criteria: **(i)** the material should be a perovskite ternary halide, **(ii)** it must be lead-free, and **(iii)** its bandgap should fall within the Shockley-Queisser limit. A high Θ_D indicates strong atomic bonds, leading to superior thermal conductivity and stability at elevated temperatures [20]. A larger bandgap provides better electronic insulation, making the material suitable for high-power and high-frequency applications. The data of Θ_D is very limited in the Materials Project database [14]. Hence, we conducted extensive data mining for halogen ternary compounds (perovskite halide), specifically involving Br, Cl, and I, leveraging the vast Materials Project database [14]. Our objective was to swiftly predict key material properties i.e. Θ_D and bandgap of lead-free halide perovskites. To achieve this, we employed the CGCNN ML model [16], which has demonstrated remarkable efficacy in predicting material properties based on crystal structures. After performing data mining, it is evident from Fig. 1 that Cl-based materials in the lead-free halide perovskite family exhibit both a large bandgap and high Θ_D. Therefore, we focused on predicting new structures within Cl-based materials in this lead-free halide perovskite family.

DFT Method
All the DFT calculations reported in this work are conducted using the Vienna Ab Initio Simulation Package (VASP) employing the projector augmented wave method (PAW) [17,18]. The exchange and correlation energies are calculated using the generalized gradient approximation parameterized by Perdew-Burke-Ernzerhof (PBE) [19]. A plane wave cut-off energy of 400 eV was utilized for all DFT calculations. The force convergence criterion for relaxing inner-atomic coordinates was set to 10^{-4} eV/ Å, and the energy convergence criterion for self-consistent DFT calculations was established at 10^{-6} eV.

A detailed flowchart illustrating our workflow is presented in Fig. 2. Our method starts initiates with data mining **(1)**, involving the thor-
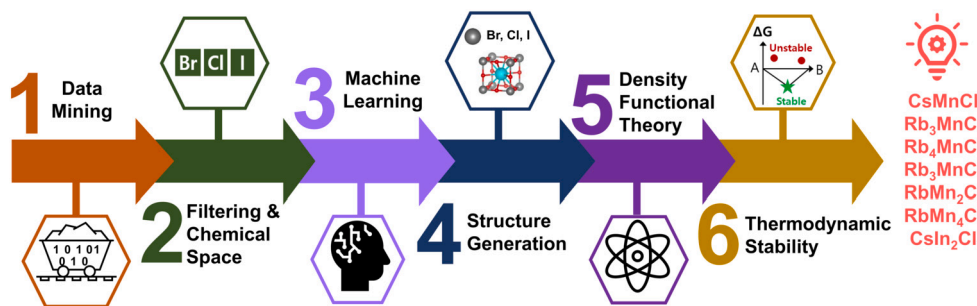
Fig. 2. Flowchart schematic outlining our methodology for discovering new materials in this work.

Table 1
The newly predicted materials underwent analysis to determine their bandgap and Debye temperature. The CGCNN model was trained using data generated from DFT-PBE calculations.

	Compound	Bandgap DFT-PBE (eV)	Bandgap CGCNN (eV)	Bandgap Type	Debye Temp. DFT-PBE (K)	Debye Temp. CGCNN (K)
1	$CsMnCl_4$	1.37	1.54	Indirect	147.23	125.57
2	$Rb_3Mn_2Cl_9$	0.00	0.23	Indirect	140.86	142.71
3	Rb_4MnCl_6	2.61	2.34	Direct	178.65	173.35
4	Rb_3MnCl_5	2.84	2.44	Indirect	158.96	160.19
5	$RbMn_2Cl_7$	0.79	0.95	Indirect	142.22	145.55
6	$RbMn_4Cl_9$	1.89	1.15	Indirect	163.49	154.51
7	$CsIn_2Cl_7$	3.01	3.18	Indirect	139.25	121.85

ough examination of extensive databases to find relevant information concerning known materials and their photovoltaic characteristics. Subsequently, the exploration and filtration of chemical space **(2)** are conducted to pinpoint possible choices based on specific criteria. The incorporation of ML technique **(3)** enhance predictive capabilities by extrapolating patterns from the accumulated data. Subsequently, the creation of crystal structures **(4)** is performed to propose novel material configurations. The utilization of density functional theory **(5)** is crucial in evaluating the electronic and optical properties of these structures, ensuring their suitability for photovoltaic application. Ultimately, thermodynamic stability analysis **(6)** is utilized to confirm the viability of the proposed materials, providing a comprehensive and rigorous strategy for anticipating new photovoltaic materials with improved precision and performance.

3. Result and discussion

Machine learning is a powerful tool for discovering new compounds, enabling the efficient analysis of vast datasets to predict material properties and identify promising candidates for novel materials [21]. There is a relationship between bandgap and Debye temperator of semiconductors [22]. Our findings unveiled intriguing trends among the studied halogen family members. Specifically, chloride-based materials exhibited a notably large bandgap and high Debye temperature compared to the bromide and iodide counterparts. This discovery provides valuable insights into the potential suitability of chloride-based halide perovskites for certain applications where these properties are advantageous. The predictive accuracy of the CGCNN model for bandgap and Debye temperature was quantified, demonstrating an impressive 83% accuracy for bandgap prediction and 90% accuracy for Debye temperature prediction. These results underscore the robustness and reliability of the CGCNN model in rapidly screening and predicting material properties (Fig. 3).

Hautier et al. [23] introduced a novel probabilistic model designed to evaluate the feasibility of ionic species substituting for one another while preserving the crystal structure. This pioneering model leverages an extensive experimental database of crystal structures for its training. Inspired by techniques from the field of machine translation, this model incorporates substitution knowledge into a specialized probability function. This function can be employed to quantitatively gauge the

probability that a particular substitution will result in the formation of another stable compound maintaining the same crystal structure. Recently, H. Kim et al. adopted a same approach, as described by Hautier et al., to predict the emergence of a novel halide perovskite, specifically Cs_3LuCl_6 [24]. This underscores the continued utility and effectiveness of the model in predicting new compounds with desired crystal structures in the realm of materials science. We provide the code used for the structure predictor, given in supplementary section(I). A similar approach has been also applied in Ref. [25] new Pr based material i.e. Pr_3AlO_6, $Pr_4Al_2O_9$, Pr_3ScO_6 and $Pr_3Sc_5O_{12}$ exhibit a large bandgap and high Debye temperature.

The Cs-Mn-Cl, Rb-Mn-Cl, and Cs-In-Cl families bandgap dataset encompasses a total number of 2608 value. The data mining is performed by using materials project [14]. In this study, we apply the equivalent ML algorithm to predict new structures in the Cs-Mn-Cl family, Rb-Mn-Cl family, and Cs-In-Cl family. The ML model identifies seven new compounds, namely **(i)** $CsMnCl_4$, **(ii)** $Rb_3Mn_2Cl_9$, **(iii)** Rb_4MnCl_6, **(iv)** Rb_3MnCl_5, **(v)** $RbMn_2Cl_7$, **(vi)** $RbMn_4Cl_9$, and **(vii)** $CsIn_2Cl_7$, all of which have zero energy above the convex hull, calculated by DFT. Detailed information about these materials is provided in Table 1 and Supplementary Section (II).

Calculating the formation energy and constructing the associated convex hull is a fundamental step in determining the energetic stability of a compound [26]. From a thermodynamic perspective, the convex hull pertains to the Gibbs free energy of the compounds at absolute zero. Our computations demonstrate the energetic stability of the newly predicted compounds, as validated by the convex hull plot presented in Fig. 4. Additional details, encompassing E_{hull} and parent atom information, can be found in Supplementary Section (III).

3.1. Bandgap and density of states

The electronic bandgap of the newly predicted compounds is computed utilizing both PBE-DFT [19] and CGCNN [16] methods. The computed bandgap values for all candidates are documented in Table 1. The Debye temperature for the newly predicted compounds was determined using CGCNN [16] and validated against DFT calculations, as detailed in Table 1. All of the predicted compounds exhibit a substantial bandgap and elevated Debye temperature. Our focus is on lead-free materials
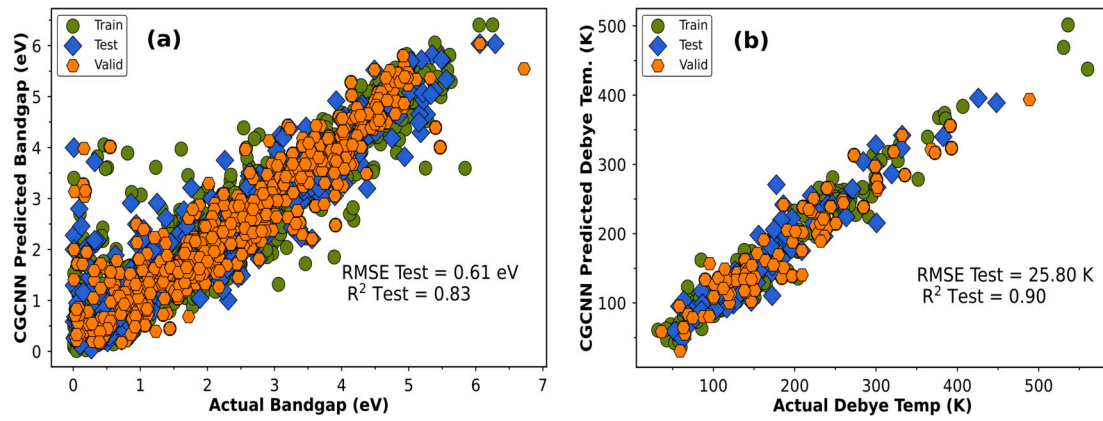
Fig. 3. The comparison plot illustrates the disparity between actual values (DFT-PBE) and CGCNN-predicted values for both **(a)** bandgap and **(b)** Debye temperature.
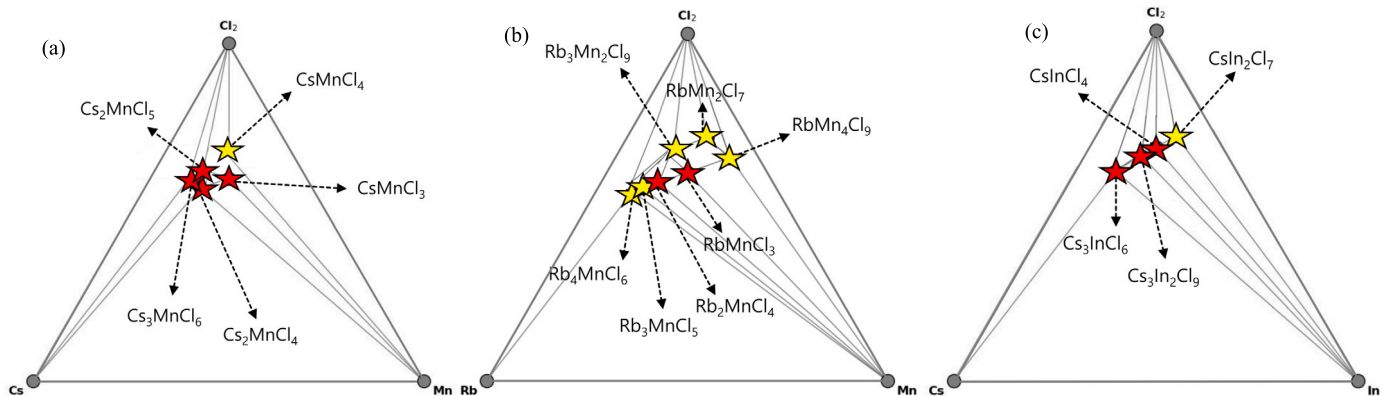

Fig. 4. The convex hull plot is shown for **(a)** the Cs-Mn-Cl family, **(b)** the Rb-Mn-Cl family, and **(c)** the Cs-In-Cl family. Known compounds from the materials-project database are marked with red stars [14], while newly predicted compounds are represented by yellow stars.
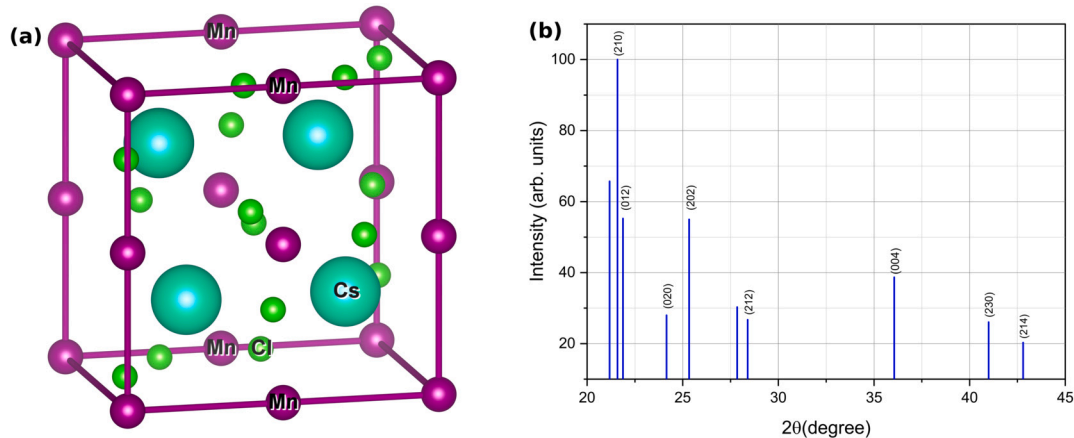

Fig. 5. $CsMnCl_4$ **(a)** crystal structure ($P2_1/c$), **(b)** simulated-XRD data generated utilizing Pymatgen's XRD module.

with bandgaps falling within the Shockley-Queisser limit [27], typically ranging from approximately 1.1 eV to 1.6 eV. Among the seven predicted compounds, only $CsMnCl_4$ possesses a bandgap of 1.37 eV, placing it within the Shockley-Queisser limit. Consequently, we delve into a detailed discussion of the material properties of $CsMnCl_4$. The structure detail and simulated x-ray diffraction spectrum [28] of $CsMnCl_4$, shown in Fig. 5.

In the context of the Shockley-Queisser limit [27], the indirect bandgap imposes limitations on the radiative recombination process, leading to lower photogenerated current and efficiency. Researchers are actively exploring innovative approaches, such as tandem structures

and advanced materials, to overcome the limitations posed by the indirect bandgap and push the efficiency of silicon solar cells closer to the Shockley-Queisser limit. This pursuit of efficiency improvements in indirect bandgap semiconductors is crucial for the continued advancement of solar cell technologies and the broader goal of sustainable energy production. Hence, the predicted material $CsMnCl_4$ consists a potential suitability for photovoltaic applications.

To gain further insights into the electronic structure of the newly predicted compound $CsMnCl_4$, we computed both the total electronic density of states (DOS) and the atomic-orbitals resolved electronic DOS, as depicted in Fig. 6**(a)**. In the anticipated compound $CsMnCl_4$, the pre-
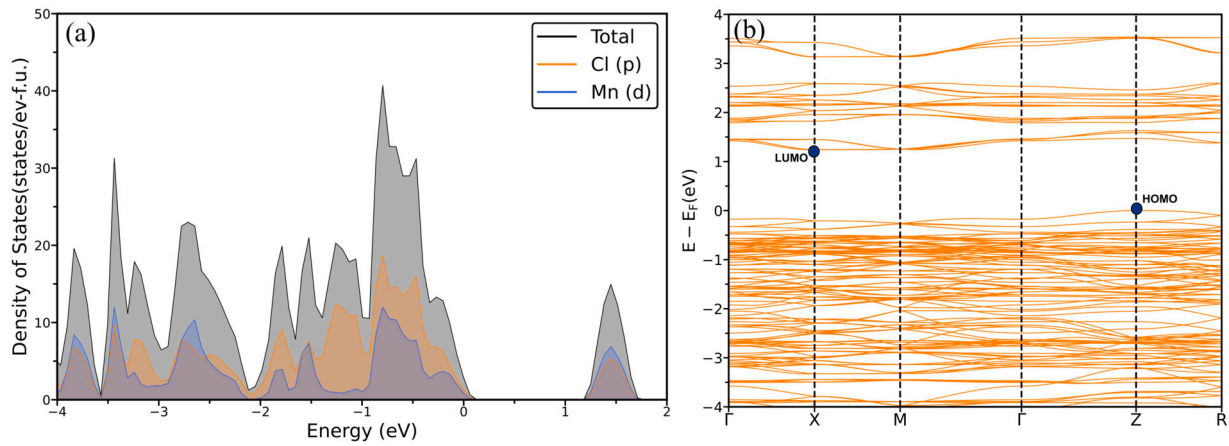
Fig. 6. The PBE calculated **(a)** DOS, **(b)** band structure, the HOMO and LUMO, showing it is an indirect bandgap.

dominant contributions arise from the *p* orbitals of chlorine and the *d* orbitals of manganese. The peak near -1 eV can be attributed to the hybridization of all constituent elements. In the context of the valence band maximum, the primary contribution to the total DOS comes from the *p* orbital of chlorine. However, for the conduction band minimum, the primary contribution to the total DOS is attributed to the *d* orbital of manganese. Hence, electronic transitions from chlorine (*p* orbitals) to manganese (*d* orbitals) are feasible.

A band structure plot visually represents the energy difference between the valence and conduction bands in a material, crucial for understanding its electrical properties and potential applications, such as semiconductor device design and solar cell efficiency optimization. To confirm the indirect bandgap behavior, we plotted the band structure of $CsMnCl_4$, as shown in Fig. 6**(b)**. This distinctly illustrates the indirect bandgap behavior of $CsMnCl_4$.

3.2. Optical properties

One can extract the linear optical properties by analyzing the frequency-dependent complex dielectric function, denoted as $\epsilon(\omega)$:

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega). \tag{1}$$

The dielectric function is expressed through its real and imaginary components, denoted as $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$, respectively, where ω signifies the photon frequency. The real component $\epsilon_1(\omega)$ is determined using the Kramers–Krönig relationship [29], while the imaginary component $\epsilon_2(\omega)$ is computed through momentum matrix elements between valence and conduction wave functions [30]. Utilizing $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$, the refractive index $n(\omega)$ and absorption coefficient $\alpha(\omega)$ can be calculated using the following formulas:

$$n(\omega) = \left[\frac{\sqrt{\epsilon_1^2 + \epsilon_2^2} + \epsilon_1}{2} \right]^{\frac{1}{2}}, \text{ and} \tag{2}$$

$$\alpha(\omega) = \sqrt{2}\omega \left[\frac{\sqrt{\epsilon_1^2 + \epsilon_2^2} - \epsilon_1}{2} \right]^{\frac{1}{2}}. \tag{3}$$

In Fig. 7, we illustrate the computed $\epsilon_1(\omega)$ and $\epsilon_2(\omega)$ as functions of ω for the anticipated compounds as per our predictions.

The real part exhibits its maximum peak around 10 eV, corresponding to $\epsilon_1(\omega) = 1.7$, in the case of $CsMnCl_4$ in Fig. 7**(a)**. The static dielectric constants can be determined by evaluating the zero-frequency limits as ω approaches 0 for $\epsilon_1(\omega)$. The static dielectric constant for $CsMnCl_4$ is found to be 1.3. In Fig. 7**(b)**, the plot of $\epsilon_2(\omega)$ indicates that

the threshold energy of the dielectric function is approximately 5 eV. This corresponds to the fundamental absorption edge, representing the optical transition between the valence band maximum (VBM) and the conduction band minimum (CBM). In the increasing energy range, the absorptive part of $\epsilon_2(\omega)$ reveals a prominent peak at approximately 15 eV. This peak is induced by the transition of Cl−3p electrons to the *s* states of cations [31].

The absorption coefficient $\alpha(\omega)$ characterizes the attenuation of light intensity as it propagates through a material over a unit distance. As shown in Fig. 8**(a)**, the onset of the absorption edge becomes noticeable at approximately 6 eV. This phenomenon is attributed to excited electrons transitioning from Cl−3p states at the upper edge of the valence band to unoccupied cation 3s states. It's important to highlight that the absorption coefficient $\alpha(\omega)$ is observed at a value below 6 eV, specifically in the ultraviolet range. Conversely, the compound $CsMnCl_4$ exhibit significant absorption, attributed to the sharp rise in the absorption coefficient beyond the absorption edge as photon energy surpasses a certain threshold. This characteristic is commonly observed in semiconductors and insulators. Fig. 8**(b)** illustrates the measured curve of $n(\omega)$ with respect to photon energy. It is noteworthy that the static refractive index $n(0)$ for incident light is 0.74 in the case of $CsMnCl_4$. Around a photon energy of 0.5 eV, the refractive index $n(\omega)$ reaches its peak value. Subsequently, the energy level gradually diminishes until reaching its minimum point, beyond which there is minimal change in the high-energy zone ($\geq$25 eV).

4. Conclusion

The seven materials newly predicted are: $CsMnCl_4$, $Rb_3Mn_2Cl_9$, Rb_4MnCl_6, Rb_3MnCl_5, $RbMn_2Cl_7$, $RbMn_4Cl_9$, and $CsIn_2Cl_7$, predicted through ML and further examined via DFT calculation. We examined the formation energy and constructed the corresponding convex hull plot, leading to the conclusion that the newly identified materials exhibit thermodynamic stability at 0 K. All of these discovered compounds demonstrated notable bandgaps and Debye temperatures, suggesting their potential usefulness in optoelectronic devices and applications involving photoluminescence. Among these, $CsMnCl_4$ stands out, with a DFT-PBE bandgap of 1.37 eV, aligning with the Shockley-Queisser limit, rendering it suitable for photovoltaic applications. Furthermore, the optical absorption and emission spectra of $CsMnCl_4$ have been calculated, further supporting their suitability for photovoltaic applications. Moreover, their high Debye temperatures further emphasize their thermal stability, positioning them as promising contenders for practical applications. It is worth mentioning that the magnetic ordering of Mn atoms may moderately change the bandgap and Debye temperature values. However, a detailed study focusing on the magnetic properties of $CsMnCl_4$ is beyond the scope of this work.

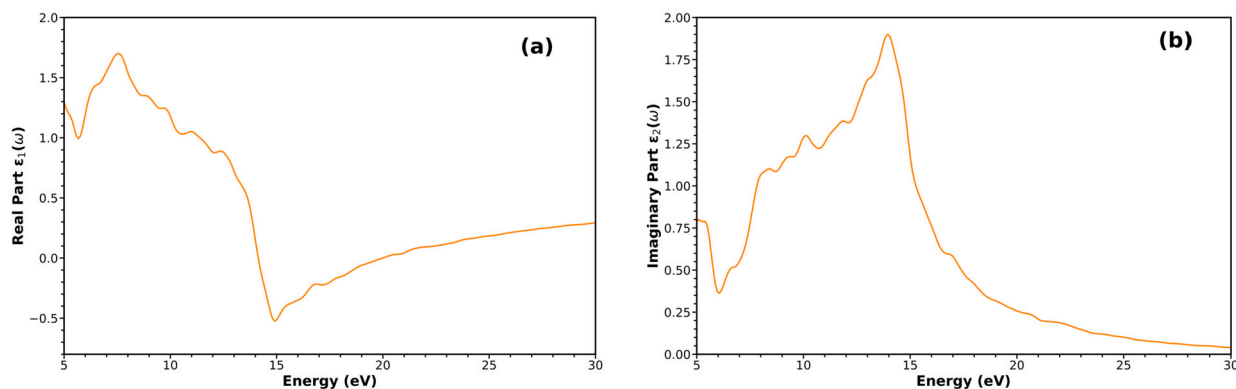**Fig. 7.** The PBE calculations yield **(a)** the real part $\epsilon_1(\omega)$ and **(b)** the imaginary part $\epsilon_2(\omega)$ of the complex dielectric function for CsMnCl$_4$.
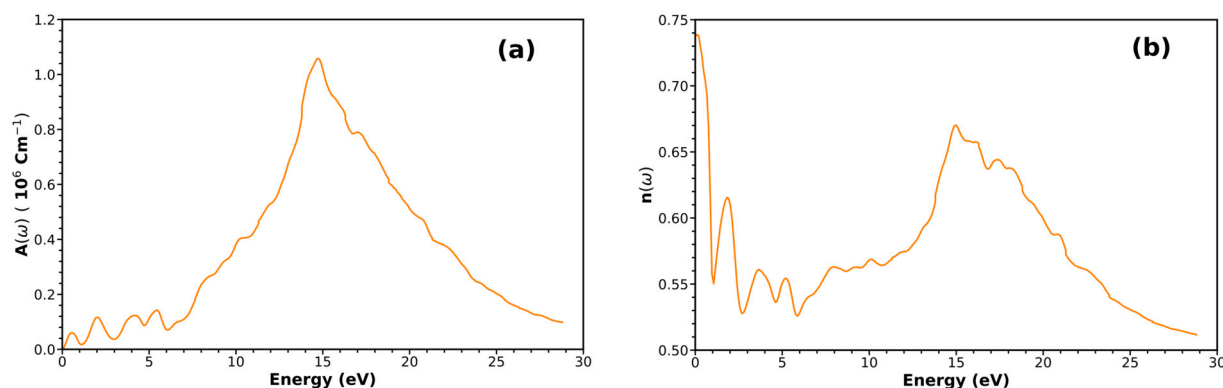

Fig. 8. The PBE calculated **(a)** absorption $A(\omega)$ and **(b)** refractive $n(\omega)$ spectrum of CsMnCl$_4$.

However, our study is not without limitations. The synthetic accessibility of these materials, particularly those with the Mn^{3+} oxidation state can be challenging due to the complex redox chemistry of Mn, as it exhibits multiple oxidation states ranging from Mn^{2+} to Mn^{7+}. Achieving the Mn^{3+} oxidation state (e.g., MnPO$_4$ [32], MnOCl [33], MnOBr [33], Mn$_2$O$_3$ [34], Mn$_2$S$_3$, Mn$_2$Se$_3$, and Mn$_2$Te$_3$ [35]) often requires precise control over reaction conditions, including temperature, pressure, and stoichiometry. Additionally, the potential for disproportionation reactions for Mn^{3+} ions and the unintended introduction of impurities during the synthesis process can lead to the formation of undesirable byproducts or instability in the synthesized compound. In summary, the synthesis of compounds with the Mn^{3+} oxidation state demands careful experimental design and optimization to overcome these challenges and achieve successful formation while maintaining stability and purity.

Looking forward, the exploration of additional compositional spaces and the incorporation of more complex machine learning models could uncover even more promising candidates for photovoltaic applications. Experimental validation of our computational predictions remains a crucial next step, which will not only test the accuracy of our computational methods but also potentially lead to the development of high-efficiency, lead-free perovskite solar cells. The recent successful discovery of Cs$_3$LuCl$_6$, a novel halide perovskite compound [24], which has been demonstrated in white-light-emitting diodes, underscores the potential of data- and ML-driven approaches in material discovery for practical applications. This achievement highlights the efficacy of combining computational predictions with experimental endeavors. Further research will likely deepen our understanding of the mechanisms underpinning the identified trends, thereby enriching our knowledge of halide perovskite materials and guiding future strategies in material design.

CRediT authorship contribution statement

The concept was conceived by Upendra Kumar and Hyeon Woo Kim. Hyeon Woo Kim, Gyanendra Kumar Maurya, Bincy Babu Raj and Sobhit Singh offered valuable suggestions, and Upendra Kumar authored the manuscript, with all authors reading and reviewing it. Ajay Kumar Kushwaha, Sung Beom Cho and Hyunseok Ko supervised the project.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cap.2024.10.012.

Data and code availability

All data and code needed to produce the work are available from the corresponding author. The data and code utilized in this study are already available in the public domain.

References

[1] W.-J. Yin, J.-H. Yang, J. Kang, Y. Yan, S.-H. Wei, Halide perovskite materials for solar cells: a theoretical review, J. Mater. Chem. A 3 (2015) 8926.

[2] E. Unger, L. Kegelmann, K. Suchan, D. Sörell, L. Korte, S. Albrecht, Roadmap and roadblocks for the band gap tunability of metal halide perovskites, J. Mater. Chem. A 5 (2017) 11401.

[3] L. Lei, Q. Dong, K. Gundogdu, F. So, Metal halide perovskites for laser applications, Adv. Funct. Mater. 31 (2021) 2010144.

[4] S. Shrestha, X. Li, H. Tsai, C.-H. Hou, H.-H. Huang, D. Ghosh, J.-J. Shyue, L. Wang, S. Tretiak, X. Ma, et al., Long carrier diffusion length in two-dimensional lead halide perovskite single crystals, Chem 8 (2022) 1107.

[5] S.B. Jathar, S.R. Rondiya, B.R. Bade, M.P. Nasane, S.V. Barma, Y.A. Jadhav, A.V. Rokade, K.B. Kore, D.S. Nilegave, A.M. Funde, et al., Facile method for synthesis of cspbbr3 perovskite at room temperature for solar cell applications, ES Mater. Manuf. 12 (2021) 72.

[6] X. Dai, Y. Deng, C.H. Van Brackle, S. Chen, P.N. Rudd, X. Xiao, Y. Lin, B. Chen, J. Huang, Scalable fabrication of efficient perovskite solar modules on flexible glass substrates, Adv. Energy Mater. 10 (2020) 1903108.

[7] A. Mathur, H. Fan, V. Maheshwari, Organolead halide perovskites beyond solar cells: self-powered devices and the associated progress and challenges, Mater. Adv. 2 (2021) 5274.

[8] W. Ning, F. Gao, Structural and functional diversity in lead-free halide perovskite materials, Adv. Mater. 31 (2019) 1900326.

[9] M. Chen, Q. Dong, F.T. Eickemeyer, Y. Liu, Z. Dai, A.D. Carl, B. Bahrami, A.H. Chowdhury, R.L. Grimm, Y. Shi, et al., High-performance lead-free solar cells based on tin-halide perovskite thin films functionalized by a divalent organic cation, ACS Energy Lett. 5 (2020) 2223.

[10] Z. Jin, Z. Zhang, J. Xiu, H. Song, T. Gatti, Z. He, A critical review on bismuth and antimony halide based perovskites and their derivatives for photovoltaic applications: recent advances and challenges, J. Mater. Chem. A 8 (2020) 16166.

[11] S.F. Hoefler, G. Trimmel, T. Rath, Progress on lead-free metal halide perovskites for photovoltaic applications: a review, Monatsh. Chem.-Chem. Mon. 148 (2017) 795.

[12] Z. Ma, L. Wang, X. Ji, X. Chen, Z. Shi, Lead-free metal halide perovskites and perovskite derivatives as an environmentally friendly emitter for light-emitting device applications, J. Phys. Chem. Lett. 11 (2020) 5517.

[13] V. Pecunia, L.G. Occhipinti, A. Chakraborty, Y. Pan, Y. Peng, Lead-free halide perovskite photovoltaics: challenges, open questions, and opportunities, APL Mater. 8 (2020) 100901.

[14] A. Jain, S.P. Ong, G. Hautier, W. Chen, W.D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, et al., Commentary: the materials project: a materials genome approach to accelerating materials innovation, APL Mater. 1 (2013) 011002.

[15] Y. Cui, M. Li, Y. Hu, Emerging interface materials for electronics thermal management: experiments, modeling, and new opportunities, J. Mater. Chem. C 8 (2020) 10568.

[16] T. Xie, J.C. Grossman, Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties, Phys. Rev. Lett. 120 (2018) 145301.

[17] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. 6 (1996) 15.

[18] G. Kresse, J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (1996) 11169.

[19] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (1996) 3865.

[20] K.A. Denault, J. Brgoch, M.W. Gaultois, A. Mikhailovsky, R. Petry, H. Winkler, S.P. DenBaars, R. Seshadri, Consequences of optimal bond valence on structural rigidity and improved luminescence properties in $Sr_xBa_{2--x}SiO_4{:}Eu^{2+}$ orthosilicate phosphors, Chem. Mater. 26 (2014) 2275.

[21] Y. Zhuo, A. Mansouri Tehrani, A.O. Oliynyk, A.C. Duke, J. Brgoch, Identifying an efficient, thermally robust inorganic phosphor host via machine learning, Nat. Commun. 9 (2018) 1.

[22] B. Ullrich, M. Bhowmick, H. Xi, Relation between debye temperature and energy band gap of semiconductors, AIP Adv. 7 (2017) 045109.

[23] G. Hautier, C. Fischer, V. Ehrlacher, A. Jain, G. Ceder, Data mined ionic substitutions for the discovery of new compounds, Inorg. Chem. 50 (2011) 656.

[24] H.W. Kim, J.H. Han, H. Ko, T. Samanta, D.G. Lee, D.W. Jeon, W. Kim, Y.-C. Chung, W.B. Im, S.B. Cho, High-throughput screening on halide perovskite derivatives and rational design of Cs_3LuCl_6, ACS Energy Lett. 8 (2023) 3621.

[25] U. Kumar, H.W. Kim, S. Singh, S.B. Cho, H. Ko, Designing pr-based advanced photoluminescent materials using machine learning and density functional theory, J. Mater. Sci. 1 (2024).

[26] G.G. Peterson, J. Brgoch, Materials discovery through machine learning formation energy, J. Phys. Energy 3 (2021) 022002.

[27] W. Shockley, H. Queisser, Detailed balance limit of efficiency of p–n junction solar cells, Renew. Energy (2018), Routledge, pp. Vol2_35–Vol2_54.

[28] H. Ling, J. Montoya, L. Hung, M. Aykol, Solving inorganic crystal structures from x-ray powder diffraction using a generative first-principles framework, Comput. Mater. Sci. 214 (2022) 111687.

[29] J.S. Toll, Causality and the dispersion relation: logical foundations, Phys. Rev. 104 (1956) 1760.

[30] H. Ehrenreich, M.H. Cohen, Self-consistent field approach to the many-electron problem, Phys. Rev. 115 (1959) 786.

[31] V. Wang, W. Xiao, D.-M. Ma, R.-J. Liu, C.-M. Yang, Structural, electronic, and optical properties of gaino3: a hybrid density functional study, J. Appl. Phys. 115 (2014) 043708.

[32] Y. Huang, J. Fang, F. Omenya, M. O'Shea, N.A. Chernova, R. Zhang, Q. Wang, N.F. Quackenbush, L.F. Piper, D.O. Scanlon, et al., Understanding the stability of mnpo 4, J. Mater. Chem. A 2 (2014) 12827.

[33] Y. Feng, Z. Wang, N. Liu, Q. Yang, First-principles prediction of two-dimensional mnox (x= cl, br) monolayers: the half-metallic multiferroics with magnetoelastic coupling, Nanoscale 15 (2023) 4546.

[34] S. Chen, F. Liu, Q. Xiang, X. Feng, G. Qiu, Synthesis of Mn_2O_3 microstructures and their energy storage ability studies, Electrochim. Acta 106 (2013) 360.

[35] B. Liu, X. Zhang, J. Xiong, X. Pang, S. Liu, Z. Yang, Q. Yu, H. Li, S. Zhu, J. Wu, Spin transport of half-metal mn2x3 with high Curie temperature: an ideal giant magnetoresistance device from electrical and thermal drives, Front. Phys. 19 (2024) 1.