FRANKFURT INTERNATIONAL CONFERENCE ON APPLIED PHYSICS AND MATERIALS SCIENCE · Vol. 1 No. 1 (2026)

Machine Learning-Accelerated Prediction of Thermal Conductivity in Layered Perovskite Structures

Felix Richter
Department of Physics, Goethe University Frankfurt, Frankfurt am Main, Germany
f.richter@example.org

Nadia Petrova
Faculty of Physics, Sofia University St. Kliment Ohridski, Sofia, Bulgaria
n.petrova@example.org

DOI: 10.5281/zenodo.0000000

Abstract *— Layered perovskite materials exhibit thermal transport properties relevant to thermoelectric and optoelectronic applications, yet first-principles calculation of thermal conductivity across large compositional spaces remains computationally expensive, limiting systematic materials screening. This study develops a machine learning framework trained on a curated dataset of density functional theory calculations to predict lattice thermal conductivity across a range of layered perovskite compositions with varying layer thickness and cation substitution. The model incorporates structural descriptors capturing octahedral distortion and interlayer spacing alongside compositional features, allowing predictions to generalize across compositional variations not present in the training set. Validation against held-out first-principles calculations shows prediction accuracy sufficient to reliably rank candidate compositions by relative thermal conductivity, substantially reducing the computational cost of preliminary screening compared to exhaustive first-principles evaluation. We apply the trained model to screen a broader candidate space, identifying several compositions predicted to exhibit notably low thermal conductivity consistent with favorable thermoelectric performance, which we recommend for targeted experimental follow-up. The approach demonstrates how machine learning surrogate models can meaningfully accelerate early-stage materials discovery workflows without eliminating the role of first-principles validation for final candidate selection.*

Index Terms *— computational materials science, machine learning, thermal conductivity, perovskite structures, materials informatics*

I. INTRODUCTION

Layered perovskite materials exhibit thermal transport properties relevant to thermoelectric and optoelectronic applications, with low lattice thermal conductivity generally favorable for thermoelectric efficiency while sufficiently high conductivity is often desired in optoelectronic device contexts [1]. First-principles calculation of thermal conductivity across large compositional spaces remains computationally expensive, since accurate phonon lifetime calculations typically require anharmonic force constant evaluation that scales unfavorably with structural complexity [2]. This computational cost limits systematic materials screening across the compositional and structural variations relevant to layered perovskite design.

Machine learning surrogate models trained on curated first-principles datasets have accelerated property prediction in other materials classes, but their application to layered perovskite thermal conductivity requires structural descriptors capable of capturing octahedral distortion and interlayer spacing effects that strongly influence phonon scattering in these systems [3]. Without descriptors specific to these structural features, generic compositional models may fail to generalize across the layer thickness and cation substitution variations of practical interest [4].

This study develops a machine learning framework trained on a curated dataset of density functional theory calculations to predict lattice thermal conductivity across a range of layered perovskite compositions.

II. RELATED WORK

Machine learning interatomic potentials and property prediction models have demonstrated substantial acceleration of materials screening workflows across oxide and halide perovskite families, generally using compositional and simple structural descriptors as model inputs. Thermal conductivity prediction specifically has received less attention than more readily calculable properties such as formation energy or band gap, in part because reliable thermal conductivity training data requires more computationally expensive anharmonic lattice dynamics calculations to generate.

III. METHODOLOGY

A curated dataset of density functional theory thermal conductivity calculations across layered perovskite compositions with varying layer thickness and cation substitution was used to train a gradient-boosted regression model incorporating structural and compositional descriptors.

A. Structural Descriptor Design and Model Training

Structural descriptors captured octahedral distortion index Δ_{oct}, interlayer spacing d, and average cation mass $\bar{m}$ alongside standard compositional features. The model was trained to minimize prediction error on held-out compositions, with performance assessed via the coefficient of determination between predicted and DFT-calculated thermal conductivity κ:

$$R^2 = 1 - \frac{\sum_i (\kappa_i^{DFT} - \kappa_i^{pred})^2}{\sum_i (\kappa_i^{DFT} - \bar{\kappa}^{DFT})^2} \tag{1}$$

Structural descriptors were retained in the final model only where their inclusion improved held-out R^2 relative to compositional features alone, confirming their contribution to generalization across unseen structural variation.

Table 1. Model validation performance versus first-principles evaluation

Approach	R^2 (held-out)	Rel. Compute Cost	Screening Throughput
Compositional features only	0.71	1× (ML)	High
Structural + compositional	0.89	1× (ML)	High
Direct DFT calculation	1.00 (ref)	∼400×	Low

IV. RESULTS

Table 1 summarizes model prediction accuracy and computational cost relative to direct first-principles evaluation across the validation compositions.

Validation against held-out first-principles calculations shows prediction accuracy sufficient to reliably rank candidate compositions by relative thermal conductivity, substantially reducing the computational cost of preliminary screening compared to exhaustive first-principles evaluation.

V. DISCUSSION

Applying the trained model to screen a broader candidate space identified several compositions predicted to exhibit notably low thermal conductivity consistent with favorable thermoelectric performance, which we recommend for targeted experimental follow-up rather than immediate adoption without further validation. The demonstrated accuracy gain from including octahedral distortion and interlayer spacing descriptors confirms that generic compositional features alone are insufficient for reliable thermal conductivity ranking in this materials class, underscoring the value of physically motivated descriptor design over purely data-driven feature selection.

VI. CONCLUSION

This study demonstrated a machine learning surrogate model capable of accelerating thermal conductivity screening across layered perovskite compositions while maintaining prediction accuracy sufficient for reliable candidate ranking. The approach illustrates how machine learning surrogate models can meaningfully accelerate early-stage materials discovery workflows without eliminating the role of first-principles validation for final candidate selection.

REFERENCES

[1] D. G. Cahill et al., "Nanoscale thermal transport," J. Appl. Phys., vol. 93, no. 2, pp. 793-818, 2003.

[2] A. Togo and I. Tanaka, "First principles phonon calculations in materials science," Scr. Mater., vol. 108, pp. 1-5, 2015.

[3] F. Richter and N. Petrova, "Machine learning descriptors for layered perovskite phonon transport," npj Comput. Mater., vol. 9, art. 112, 2023.

[4] R. Ramprasad et al., "Machine learning in materials informatics: Recent applications and prospects," npj Comput. Mater., vol. 3, art. 54, 2017.

[5] J. Carrete et al., "Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling," Phys. Rev. X, vol. 4, art. 011019, 2014.

[6] G. K. H. Madsen and D. J. Singh, "BoltzTraP: A code for calculating band-structure dependent quantities," Comput. Phys. Commun., vol. 175, no. 1, pp. 67-71, 2006.