

報道機関 各位

国立大学法人東北大学

AI を活用し 膨大な実験と理論計算のデータを統合した 材料マップを開発

—— 新しい機能性材料開発期間の大幅短縮に期待 ——

【発表のポイント】

- 人工知能（AI）を実現する一手法の機械学習を活用して実験データと計算データを融合し、材料の構造類似性をひと目で把握できる「材料マップ」^{（注1）}を構築しました。
- 未知の高性能材料の類似体を迅速に抽出でき、研究者が既存の合成手法を転用して次のターゲット材料を容易に選定することが可能になります。
- 今後、新材料開発期間の大幅な短縮につながることを期待されます。

【概要】

これまで材料開発では、理論計算による予測と実験による検証が分離しており、効率的な材料探索が課題となっていました。

東北大学学際科学フロンティア研究所寄附研究部門「ナノ材料プロセスデータ科学」の橋本佑介特任准教授と筈居高明教授、同大学材料科学高等研究所（WPI-AIMR）の賈雪（ジャ シュエ）助教と李昊（ハオ リー）教授らの研究グループは、実験と代表的な理論計算の第一原理計算データを統合した大規模データセットにメッセージパッシング型グラフニューラルネットワーク（MPNN）^{（注2）}を適用し、材料の熱電特性と構造類似性を一望できる「材料マップ」の構築に成功しました。マップ上では構造的に近い材料が近接して配置されます。似た構造の材料は似た手法で合成・評価されることが多く、既存の合成手法を生かして次のターゲット材料を選定するときの効率化に貢献します。

このマップを活用することで、実験研究者が未知の高性能材料において構造的に類似するものを短時間で抽出し、既存の合成手法を転用して次のターゲット材料を容易に選定することが可能になります。新材料開発期間の大幅な短縮につながることを期待される成果です。

本研究成果は、2025 年 7 月 28 日（現地時間）に科学誌 APL Machine Learning のオンライン版にて公開されました。

【詳細な説明】

研究の背景

無数に存在する材料の中から目的に適した最良のものを選択することは、容易ではありません。材料のビッグデータとデータ科学的手法を融合させた新しい研究分野であるマテリアルズインフォマティクス（MI）は、この課題に対し現実的な解決策を提示します。MI を活用して効率的な材料開発を実現するには、材料に関する大量のデータを収集する必要があります。

材料特性のデータは主に計算と実験から得られます。計算データは米国政府主導のデータ駆動型材料開発（Materials Genome Initiative: MGI）を契機に集積が進み、「Materials Project」や「AFLOW」などの大規模計算データベースが整備されています。一方、実験データは論文やレポートに散在していますが、国立研究開発法人物質・材料研究機構（NIMS）が文献由来の実験データを収集した Web システム「StarryData2」として公開しています。

そこで本研究チームは、StarryData2 から抽出した熱電材料^{（注 3）}のデータを用いて高精度（ $R^2 = 0.90$ ）の機械学習モデルを構築し、そのモデルを Materials Project に登録された 1,114 件の計算データに適用して、計算・実験を統合したデータセットを作成しました。さらに、その解析結果から未報告の有望な熱電材料としてゲルマニウム・テルル・ヒ素化合物の $\text{Ge}_2\text{Te}_5\text{As}_2$ と $\text{Ge}_3(\text{Te}_3\text{As})_2$ を提案し、2024 年に論文として報告しました（参考文献 1）。

今回の取り組み

本研究では、前回作成したデータセットに対して材料のグラフ表現と高度な機械学習の深層学習を組み合わせた高度なデータ科学手法を適用し、特徴を反映した材料マップを作成しました（図 1）。具体的には、結晶構造をグラフとして表現し、メッセージパッシング型グラフニューラルネットワーク（MPNN）によって高次元ベクトルへ変換した後、深層学習を用いて材料特性の予測モデルを構築しました。この過程で得られたベクトル表現を次元削減手法（t-SNE^{（注 4）}）で 2 次元に圧縮し、材料マップを生成しました（図 2）。解析の結果、MPNN に基づくグラフ表現が材料の熱電特性の優位性と構造の複雑性を捉え、明瞭な構造を持つ材料マップを生み出すことを確認しました。

似た構造を持つ材料は、似た手法・装置・プロセスで合成・評価されるため、材料科学では従来から構造類似性を手掛かりに材料を体系化し、個別に詳細な探索を行ってきました。本研究が実現した、材料特性と構造類似性を同時に可視化する材料マップは、多数の材料を俯瞰的に把握し、より効果的かつ効率的な材料探索を可能にする基盤になると期待されます。

今後の展開

今回開発した材料マップは、実験研究者が構造的に類似した材料を短時間で

抽出し、合成条件を迅速に設定できるプラットフォームとして機能します。今後は熱電材料にとどまらず、磁性材料や表面と内側で電気的特性が異なるトポロジカル材料など多様な機能性材料へ適用範囲を拡大し、さらに磁性・化学的性質・トポロジカル特性などの追加物性情報を統合することで、包括的な材料設計支援システムの構築を目指します。これにより、社会実装までのリードタイムを大幅に短縮し、革新的材料の早期実用化へと貢献することが期待されます。

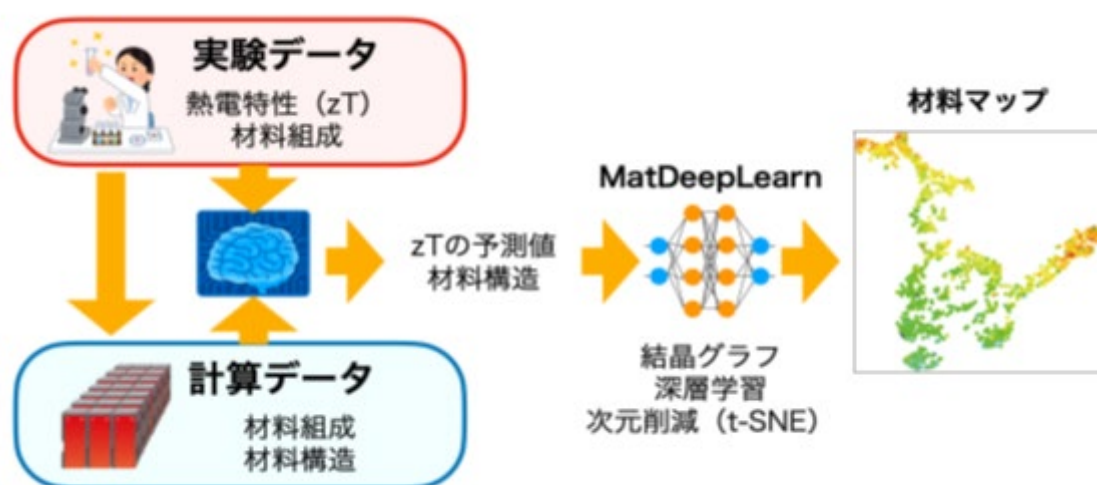


図 1. 本研究におけるデータ分析の流れ

実験データと計算データを機械学習を通して融合し、結晶グラフ、深層学習、次元削減などのデータ科学的手法を駆使して、材料マップを構築しました。

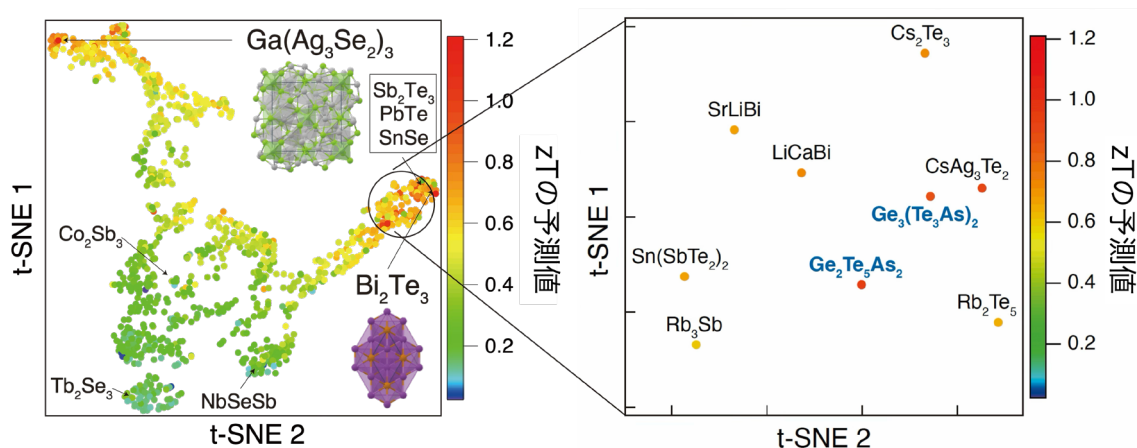


図 2. 開発した材料マップ（左）とその拡大図（右）

材料の熱電特性 (zT) と材料の構造類似性を同時に反映する材料マップにより、これらが同時に類似する材料の効率的な探索を実現します。

【謝辞】

本研究は、東北大学学際科学フロンティア研究所寄附研究部門「ナノ材料プロセスデータ科学」の支援を受けて実施されました。また、日本学術振興会 科学研究費助成事業（課題番号 JP25H01508）からの資金提供を受けています。さらに、東北大学 AIMR の ATP プロジェクトおよび Yong P. Chen 教授研究グループによる計算支援を受けました。

【用語説明】

注1. 材料マップ

高次元材料データを 2 次元平面上に投影し、材料間の類似性を直感的に可視化したもの。

注2. メッセージパッシング型グラフニューラルネットワーク (MPNN)

原子をノード、結合をエッジとして表現したグラフ上で、各ノードが隣接ノード・エッジからメッセージを受け取り集約・更新し、材料物性を予測する汎用的グラフニューラルネットワーク手法。

注3. 熱電材料

温度差を電圧に変換する材料で、廃熱回収などのエネルギー利用に期待。

注4. t-SNE (t-distributed Stochastic Neighbor Embedding)

局所的類似性を保ちながら大域的構造も可視化する非線形次元削減手法。

【参考文献】

1. Jia, X., Aziz, A., Hashimoto, Y., and Hao L. Sci. China Mater. (2024).
<https://doi.org/10.1007/s40843-023-2777-2>

【論文情報】

タイトル : A Materials Map Integrating Experimental and Computational Data via Graph-Based Machine Learning for Enhanced Materials Discovery

著者 : Y. Hashimoto*, X. Jia, H. Li, T. Tomai

*責任著者 : 東北大学 学際科学フロンティア研究所 特任准教授 橋本 佑介

掲載誌 : APL Machine Learning

DOI: <https://doi.org/10.1063/5.0274812>

【問い合わせ先】

（研究に関すること）

東北大学学際科学フロンティア研究所

特任准教授 橋本 佑介

TEL: 022-795-5630

Email: yusuke.hashimoto.b8@tohoku.ac.jp

（報道に関すること）

東北大学材料科学高等研究所（WPI-AIMR）

広報戦略室

TEL: 022-217-6146

Email: aimr-outreach@grp.tohoku.ac.jp