

2026年6月15日

東京科学大学

神奈川県立産業技術総合研究所、東北大学

AIの予測根拠を解読して材料設計指針を導く新手法を開発 －マテリアルズ・インフォマティクスによる材料探索の加速－

【ポイント】

- 物性予測 AI の内部情報から材料設計の指針を抽出する汎用的な新手法を開発
- 2,681 種の物質の光吸収スペクトルを学習した AI に適用して手法の有効性を実証
- 物性の支配因子となる元素と局所構造を自動で特定

【概要】

東京科学大学 (Science Tokyo) 総合研究院 フロンティア材料研究所の高橋亮助教、大場史康教授 (神奈川県立産業技術総合研究所 (KISTEC) プロジェクトリーダー兼任)、同大学 物質理工学院 材料系の高松新修士課程学生 (研究当時) は、東北大学 金属材料研究所の熊谷悠教授と共同で、材料のスペクトルデータを予測する人工知能 (AI) モデルを解釈し、類似する材料をグループ化することで、材料設計の指針となる情報を抽出する新手法を開発しました。

材料科学において、所望の材料特性の発現に関わる構成元素や原子配列の特徴を見いだすことは、材料設計指針の構築や物性発現機構の解明のために重要です。一方で近年、AI によって材料の物性を高い精度で予測できるようになってきており、そうした予測モデルの内部には材料設計に有用な知見が含まれていると期待されます。しかし、特にスペクトルデータについては、その知見を引き出すための解釈手法は確立していませんでした。

本研究では、結晶構造から物性を予測する AI モデルから材料ごとの特徴を抽出し、類似した特徴とスペクトル形状を持つ材料をグループ化する手法を提案しました。**第一原理計算** (用語 1) に基づく 2,681 種の金属酸化物や**カルコゲナイド** (用語 2) 等の光吸収スペクトルデータベースを構築して AI モデルの学習に用い、本手法を適用した結果、スペクトル特性と構成元素・**配位環境** (用語 3) の両方が類似した材料のグループに分類できました。さらに、この材料グループ内の共通点やグループ間の相違点を考察することで、スペクトル特性を決定づける元素種や局所構造を明らかにすることに成功しました。

本手法は、光吸収スペクトルに限らずさまざまな物性データに適用可能であり、**マテリアルズ・インフォマティクス** (用語 4) を用いた材料設計指針の獲得に広く貢献することが期待されます。

本成果は 6 月 15 日付の「*Advanced Intelligent Discovery*」誌に掲載されました。

●背景

材料科学分野において、所望の物性を持つ有望な材料群を見出し、その物性の鍵となる構成元素や原子配列の特徴を明らかにすることは、材料設計指針を構築するために重要です。近年、機械学習の発展により、バンドギャップや融点のような、1つの数値で表される物性の予測やその解釈手法は大きく進歩してきました。例えば本研究グループでは、機械学習の予測モデルに基づいて、対象とする物性に応じた無機材料の自動分類と設計指針の抽出が可能であることを示しています（参考文献1）。

一方で、光吸収スペクトルや誘電関数のようなスペクトルデータについても、深層学習モデルによって高い精度で予測できることが示されてきています。このような予測モデルの内部には、スペクトル特性と構成元素・原子配列の関係に関する、材料設計に有用な知見が含まれていると期待されます。しかし、1つの数値で表される物性とは異なり、スペクトルデータの予測モデルを解釈して材料設計指針を抽出することは困難でした。

そこで本研究では、**グラフニューラルネットワーク（GNN、用語5）**の内部情報に基づく特徴抽出と、**階層的クラスタリング（用語6）**を組み合わせることで、スペクトル予測モデルから材料設計の指針を引き出す手法の開発に取り組みました。

●研究成果

本研究ではまず、手法の検証対象として、光吸収スペクトルのデータベースを構築しました。光吸収スペクトルは太陽電池、光触媒、光学コーティングなどの性能に直結する重要な物性です。**Materials Project データベース（用語7）**から取得した金属酸化物・カルコゲナイド・関連化合物に対してハイスループット第一原理計算を実施し、2,681種の物質における光吸収スペクトルのデータベースを得ました。

次に、上述の計算データベースを用いて、GNNの一種である**ALIGNN（用語8）**による予測モデルを構築し、この予測モデルの解釈を試みました（図1）。ALIGNNは元素種や結合距離・角度をもとに物性を予測します。そのため、訓練済みALIGNNモデルの中間層から材料ごとの特徴を抽出し、クラスタリングを適用して材料を分類すれば、スペクトル形状が類似していて、さらに起源となる構成元素・配位環境が共通した物質群に分類できます。

この手法で、実際に2,681種の物質を96グループに分類した結果、各グループ内でスペクトル形状が類似していることが確認できました（図2）。さらにスペクトル形状だけでなく、構成元素や配位環境などにも明確な共通性が見いだされ、グループ内の共通点やグループ間の相違点を考察した結果、スペクトル形状を支配する因子を抽出することができました。

結果の一例として、バナジウム (V^{5+}) やクロム (Cr^{6+}) を含む酸化物群を見ると V^{5+} や Cr^{6+} を四面体配位で含み、アルカリ金属に囲まれた材料群（クラスター74）と、同じく V^{5+} や Cr^{6+} を含みつつも異なる配位環境を持ち、吸収端のエネルギーがクラスター74

と比較して小さい材料群(クラスター70)が本手法により明確に区別されました(図3)。

また、本研究の予測モデルは陽に電子構造を学習していませんが、アルカリ金属がなぜ吸収端エネルギーの違いをもたらすかを考察するために、第一原理計算で得られた電子構造を解析しました。クラスター74 に属する材料ではアルカリ金属が四面体ユニットを孤立させる役割を果たしており、**伝導帯**(用語9) 下端の電子状態が四面体ユニット周辺に局在しやすくなっています。一方、クラスター70 に属する材料ではそのような孤立構造が形成されにくく、伝導帯下端の電子状態が結晶全体に広がる傾向があります。このように材料のグループ間の相違点として見いだされたアルカリ金属が、スペクトルの立ち上がり位置の差の原因であることが裏付けられました。

注目すべきは、配位環境がスペクトルにどう影響するかといった化学的知識を AI モデルに一切与えていないにもかかわらず、このような化学的に意味のある分類が自動的に実現された点です。これは、AI モデルが結晶構造と物性データのみから電子的な特徴を学び取っていることを示しています。

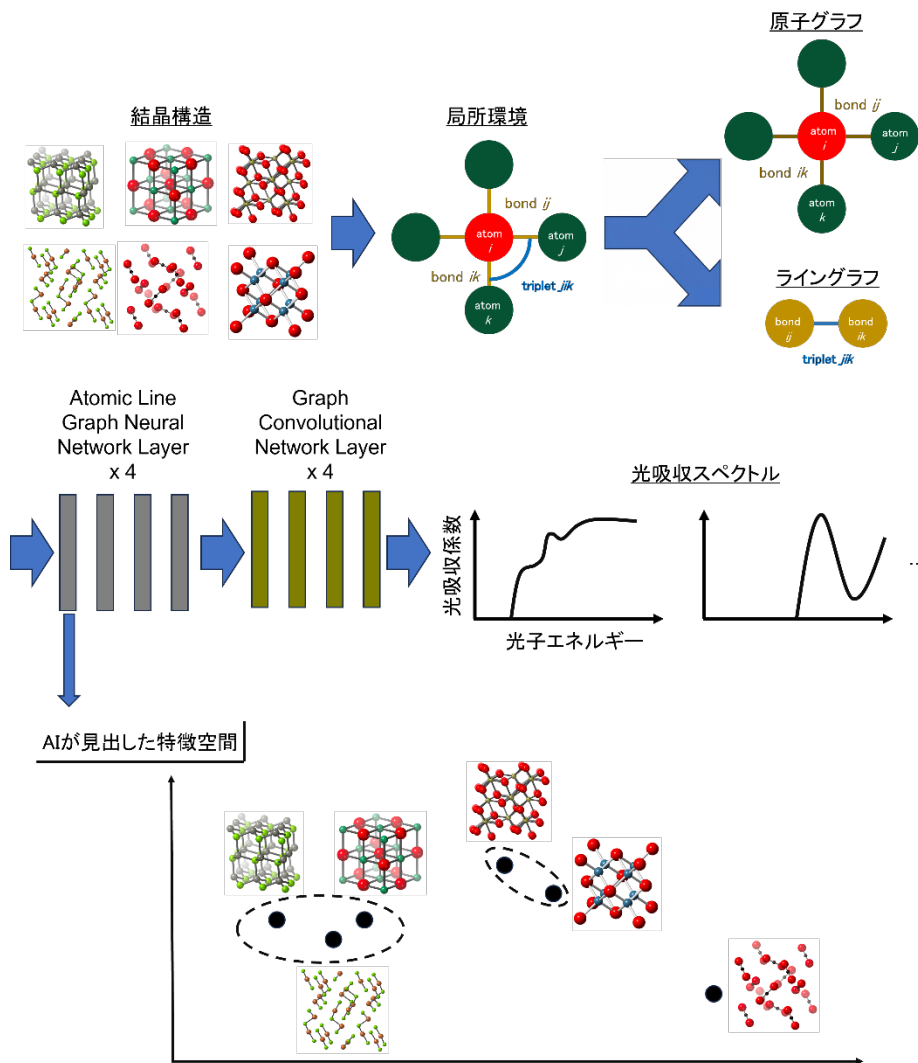


図1 本研究で提案した手法の概略図



図2 深層学習に基づいて96グループに分類した物質の光吸収スペクトル。スペクトルデータの直接的な分類でないにもかかわらず、各グループ内でスペクトル形状が類似している。

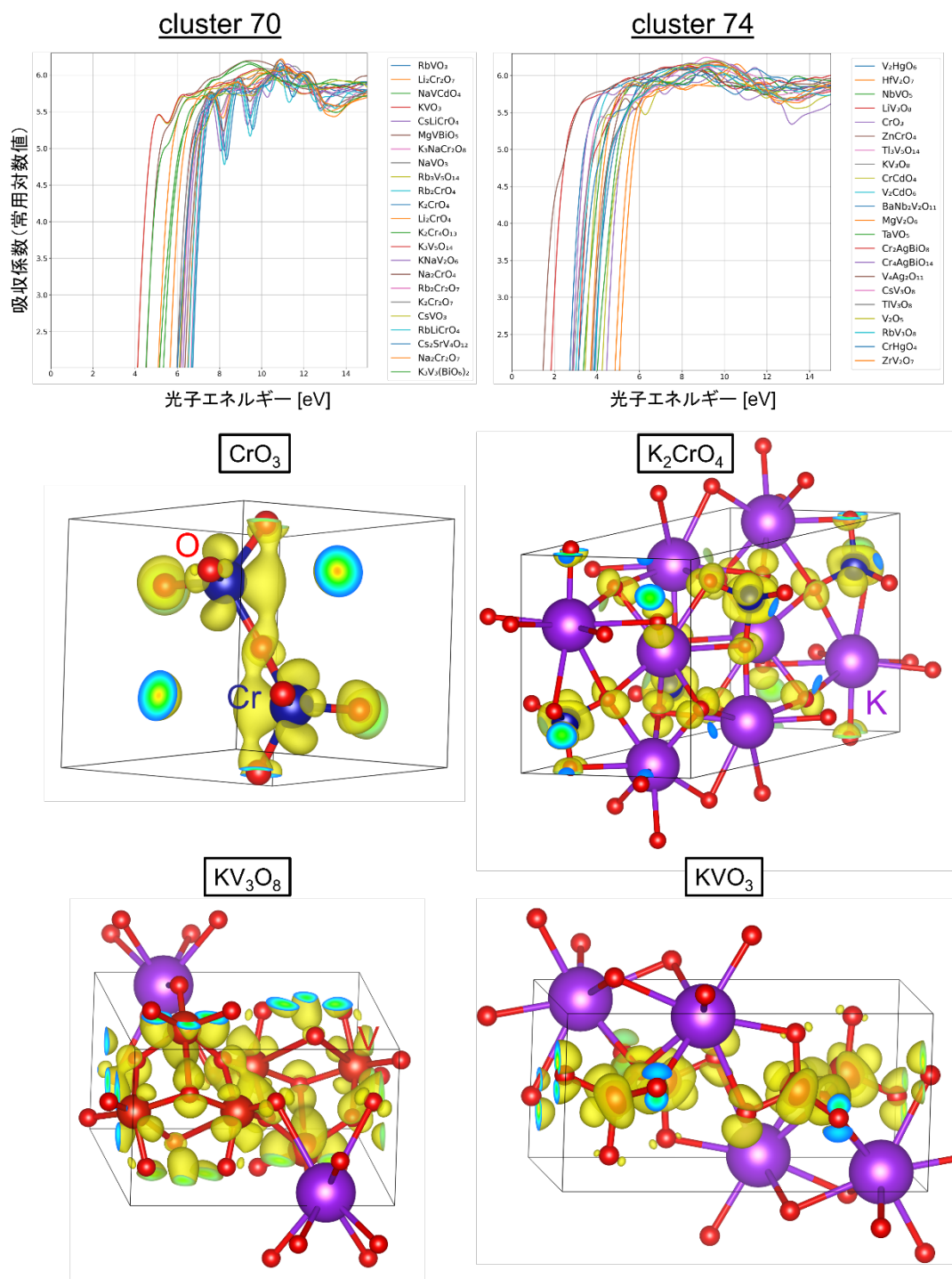


図3 バナジウムまたはクロムを含む物質群（クラスター74 とクラスター70）の分類結果と代表的な結晶構造。（上）それぞれの物質群に含まれる物質の吸収スペクトル、（下）代表的な物質の結晶構造と伝導帯下端の電子状態。

●社会的インパクト

近年、材料科学のさまざまな分野で、ハイスループット計算や自動実験技術の発展によって、大規模な材料データベースが構築されています。しかし、そうした材料データベースを機械学習に用いた高精度な予測モデルの構築が可能であっても、特にスペクトルデータについては予測根拠や材料設計に関する知見を抽出することは困難でした。本手法を用いて、深層学習モデルの予測根拠から材料設計の指針が示されれば、材料科学的な知識を人間が理解可能な形で獲得することが明確・容易になり、さまざまな材料の研究・開発が加速すると期待されます。

●今後の展開

本研究グループは、約 10 万種の物質の電子・光学特性を評価した大規模な計算材料データベースを保有しており、今後は本手法をこのデータベースに適用することで、半導体材料や光学材料の設計指針を提案していきます。さらに本手法は、ALIGNN 以外の GNN アーキテクチャや実験データで訓練されたモデルにも原理的には適用可能であり、誘電関数、電子状態密度、フォノンスペクトルなど光吸収スペクトル以外のさまざまな物性データの解釈に展開することが可能です。

●付記

本研究は日本学術振興会 科学研究費助成事業 (JP24K08562、JP24H00376)、文部科学省 データ創出・活用型マテリアル研究開発プロジェクト事業 (JPMXP1122683430)、国際・産学連携インヴァースイノベーション材料創出プロジェクト拠点 DEJI²MA プロジェクト、KISTEC 脱炭素化対策事業の助成を受けて行われました。計算には東京科学大学のスーパーコンピュータ TSUBAME4.0 を使用しました。

【参考文献】

- [1] N. Sato, A. Takahashi, S. Kiyohara, K. Terayama, R. Tamura, and F. Oba, "Target Material Property-Dependent Cluster Analysis of Inorganic Compounds", *Advanced Intelligent Systems*, 2024, 6, 2400253.

【用語説明】

- (1) **第一原理計算**: 量子力学の基本原理に基づいた理論計算。物質の電子構造やエネルギーを計算することにより、電子・光学特性や安定性、構造など多様な物性を予測することができる。
- (2) **カルコゲナイド**: 酸素以外の硫黄 (S)、セレン (Se)、テルル (Te) などの第 16 族元素 (カルコゲン) と金属元素の化合物の総称。
- (3) **配位環境**: ある原子の周囲に位置する隣接原子の数や配置のこと。例えば正四面体配位では、中心原子の周りに 4 個の原子が正四面体の頂点に位置する。

- (4) **マテリアルズ・インフォマティクス**: 材料科学における実験および理論計算の結果に対して、機械学習などのデータ科学手法を適用することで、膨大な種類の材料やその性質を俯瞰的に扱うアプローチ。
- (5) **グラフニューラルネットワーク (GNN)**: 原子をノード (点)、原子間結合をエッジ (線) とするグラフ構造を入力として、物質の特性を予測する深層学習の手法。
- (6) **階層的クラスタリング**: 類似した特徴を持つデータ同士をグループ (クラスター) にまとめる手法の一つ。データ間の類似度に基づき、最も似たもの同士から順に統合していくことで樹形図 (デンドログラム) を構成し、任意の粒度での分類が可能。
- (7) **Materials Project データベース**: 材料科学分野の大規模オープンデータベース。第一原理計算により得られた物質の構造や特性に関するさまざまな情報を提供している。
- (8) **ALIGNN (Atomistic Line Graph Neural Network)**: GNN の一種であり、結晶構造に含まれる原子間の結合距離や結合角度の情報から物性を予測する。
- (9) **伝導帯**: 半導体や絶縁体では電子に占有されていないエネルギー帯。その中でエネルギーの低い状態は光吸収特性に特に大きく影響する。

【論文情報】

掲載誌: *Advanced Intelligent Discovery*

論文タイトル: Deep Learning-Based Extraction of Promising Material Groups and Common Features from High-Dimensional Data: A Case of Optical Spectra of Inorganic Crystals (深層学習に基づく高次元データからの有望材料群およびその共通の特徴の抽出: 無機結晶の光学スペクトルの事例)

著者: Akira Takahashi, Yu Kumagai, Arata Takamatsu, and Fumiyasu Oba (高橋亮、熊谷悠、高松新、大場史康)

DOI: 10.1002/aidi.202600007

【研究者プロフィール】

高橋 亮 (タカハシ アキラ) Akira Takahashi

東京科学大学 総合研究院 フロンティア材料研究所 助教

研究分野: マテリアルズ・インフォマティクス

大場 史康 (オオバ フミヤス) Fumiyasu Oba

東京科学大学 総合研究院 フロンティア材料研究所 教授

神奈川県立産業技術総合研究所 (KISTEC) プロジェクトリーダー (兼任)

研究分野: 計算材料科学

【お問い合わせ先】

(研究に関すること)

東京科学大学 総合研究院 フロンティア材料研究所 助教

高橋 亮

Email : takahashi.a.f9db@m.isct.ac.jp

(報道取材申し込み先)

東京科学大学 総務企画部 広報課

取材申し込みページ : <https://www.isct.ac.jp/ja/001/media>

Email : media@adm.isct.ac.jp

TEL : 03-5734-2975 FAX : 03-5734-3661

神奈川県立産業技術総合研究所 研究開発部 研究推進課

Email : rep-kenkyu@kistec.jp

TEL : 044-819-2034 FAX : 044-819-2026

東北大学 金属材料研究所 情報企画室広報班

Email : press.imr@grp.tohoku.ac.jp

TEL : 022-215-2144 FAX : 022-215-2482