

研究報告書

「材料計測データのモダリティ変換」

研究期間：2020年4月～2022年3月
研究者番号：50245
研究者：鈴木 雄太

1. 研究のねらい

現在の材料開発では、日々取得される計測データの効率的な解析が課題となっています。本研究では、計測データの解析、すなわち計測データから材料の特徴を推定する操作をデータのモダリティ変換として捉え、機械学習の枠組みに乗せることを着想しました。これにより、データ解析において手間と時間を要していた試行錯誤が不要となり、多種多様なデータを自動かつ高速に解析できるようになると見込まれます。さらに、物理モデルに基づく既存のデータ解析法を自動化し、機械学習モデルによる予測と組み合わせることにより、単に高速だけでなく、物理的な妥当性を担保した手法とすることを目指します。これら一連のアイデアを、材料の結晶構造(原子の3次元配置)の分析法として普及している粉末X線回折(XRD)法において実証します。

2. 研究成果

(1)概要

物質の性質は結晶構造によって支配されるため、結晶構造の同定は物質科学の始点です。結晶構造を調べるために、粉末X線回折(XRD)法という測定法が広く用いられていますが、測定したXRDパターンから結晶構造の情報を得るためには、条件を変えながらシミュレーションを繰り返す必要があり、このデータ解析が研究のボトルネックになっていました。そこで本研究では、シミュレーションを使わず機械学習でXRDパターンから結晶構造(3次元の原子配置)を直接予測するアプローチと、数理最適化の技術を応用して既存のシミュレーション法を自動化するアプローチの2つで研究に取り組みました。これら2つは相補的なものであり、2つを組み合わせることで、「XRDから結晶構造を高速に予測→予測した構造をベースにシミュレーションし、さらに精密な構造を得る」というワークフローが実現します。アプローチ1については、ディープニューラルネットワーク(DNN)を用いて任意の結晶構造を出力する手法が確立されていないことが課題となっていました。他のACT-I研究者の皆様との共同研究においてこれを開発し、DNNを用いてXRDパターンから結晶構造を直接予測することを可能にしました。さらにその過程において、DNNを用いて結晶構造のベクトル表現を与える手法を考案しました。これは結晶構造の持つ抽象的な特徴を計算機上で扱えるようにする基礎技術であり、情報の検索や可視化など様々な応用が期待されます。アプローチ2については、シミュレーションに用いる物理モデルのパラメータ調整を、数理最適化の一分野であるブラックボックス最適化の問題として扱うことにより、シミュレーションによるXRDのデータ解析を自動化することができました。これにより熟練者が手作業で数時間以上かかっていたデータ解析を自動かつ30分程度で終わるようになり、熟練者と同等以上の精度が得られることも実証できました。

(2) 詳細

研究テーマ 1「XRD パターンから結晶構造を直接予測する機械学習技術の確立」

XRD パターンから結晶構造(3次元の原子配置)を直接予測するためには、結晶構造を何らかの方法で機械学習モデルから出力する必要がありますが、結晶構造を出力する手法は確立されておらず、この技術要素の開発が必要でした。結晶構造の出力における課題は、結晶構造を構成する原子の個数が不定であり、画像のように一定サイズのデータで表現するのが難しいことにあります。そこで、原子の数や位置を直接学習するのではなく、「原子が配列している場(field)」を学習するというアプローチによって、任意の原子個数や形の結晶構造の出力を可能にしました。このアプローチに基づく結晶構造デコーダーを DNN により構成し、畳み込みニューラルネットワークに基づく XRD エンコーダーと組み合わせることで、入力された XRD パターンから結晶構造を予測し直接出力することを可能にしました(論文準備中)。

研究テーマ 2「計測データへの物理モデルのフィッティングを自動化する技術の開発」

一般的に、計測データから材料の情報(物理量)を得るためには、様々な条件でシミュレーションを行い、計測データに最も一致するシミュレーション条件がその計測データに対応した物理量である、というアプローチを取ります。よって、シミュレーション条件の試行錯誤のために大きな手間と時間に加え熟練が必要とされていました。本研究では、この問題が機械学習モデルのハイパーパラメータ調整(HPO)と同じであると気づき、HPO と同じ枠組みで解決できるのではと着想しました。このアイデアのもと、シミュレーションに用いる物理モデルのパラメータ調整を、数理最適化の一分野であるブラックボックス最適化の問題として定式化し、ベイズ最適化に基づきこの最適化問題を解くことによって、シミュレーションによる XRD のデータ解析を自動化することができました。これにより熟練者が手作業で数時間以上かかっていたデータ解析を自動かつ 30 分程度で終わるようになり、熟練者と同等以上の精度が得られることも実証できました。さらに、自動化により多くの設定を探索できるようになったことで、手動探索では見つけられなかった構造の候補も発見できることを示しました。

2021 年 11 月現在、上記の 2 つのテーマを統合した、end-to-end な自動 XRD データ解析の実証を進めています。

3. 今後の展開

まず材料計測に関して、XRD の自動データ解析については、ACT-I 本フェーズおよび加速フェーズの研究においてある程度道筋をつけることができたと考えています。開発した手法について、ユーザーからのフィードバックを受けながら改善点の洗い出しと社会実装を進めていきます。

これらの研究で提案した方法論およびアイデアは XRD に限らず他の計測手法にも応用可能であることから、他の研究グループと連携しつつ応用の幅を広げていこうと考えています。特に、

DNN を使ったモダリティ変換の枠組みでは、DNN への入力に XRD 以外の計測モダリティ(例:X線吸収スペクトル、中性子回折パターン、電子顕微鏡画像など)を増やすことで、様々な情報を統合した学習と予測が可能になります。材料開発においては、マイクロ・メゾスコピック・マクロと、材料の空間的な階層性を統合して材料の性質を理解・モデリングすることが重要な研究テーマとなっていることから、各特徴に対応した計測モダリティを組み合わせたマルチモーダル学習によってこの問題に貢献できると考え、ACT-I 研究を発展させる形で研究に取り組んでいきます。

さらに、材料の計測を離れた展開も計画しています。日本刀が焼入れにより硬さを増すように、材料は同じ成分でも合成プロセス(=レシピ)によってその性質が大きく変わります。よって、狙った特性を備える材料の合成プロセスの開発は材料開発における中心的な問題の一つですが、プロセス開発におけるデータや情報処理技術の活用は始まったばかりです。今後、これまでの研究の知見を活かして、プロセスの記述や予測、生成に取り組んでいきます。

上記のテーマを中心として、最新の情報処理技術や機械学習の技術を応用することによる材料の問題解決に取り組むつつ、機械学習やデータ分析分野にも新たな課題を提供することを目指して、分野間をつなぐ研究を深めていきたいと思っております。

4. 自己評価

- 研究目的の達成状況について
 - 研究目的は概ね予定通り達成することができたと考えている。また研究目的に向かう過程で、結晶構造の embedding 学習など副次的な成果を得ることもできた。一方で、査読に遅れが生じるなどの背景はあったものの、論文の出版に時間を要している点は反省点である。
- 研究の進め方について
 - COVID-19 の影響で学会参加や海外の研究室との連携が難しくなるといった影響が出たが、研究費は計算リソースのほか論文投稿費等として有効に活用することができた。
 - 研究実施体制も適正であったと考えている。
- 研究成果の波及効果
 - 本研究で開発した手法はオープンソースソフトウェアとして公開しており、既に国内外の大学や研究機関、企業への普及が進みつつあることから、実用という観点での波及は着実に進んでいる。
 - これから発表予定の研究成果も含め、本研究は様々な材料に適用可能で、材料開発を支える基盤技術の一つとなるものである。よって高性能な材料の開発を加速するという形で、本研究成果は間接的にも科学技術や産業、社会に波及していくと見込まれる。
- 研究課題の独創性・挑戦性
 - 加速フェーズの研究課題は、物理法則からのボトムアップを離れ、いわばトップダウンな考え方に基いて計測データの解析を機械学習のモダリティ変換とみなすという点で、材料科学的には非常に独創性の高いものであったと考える。技術的にも、結晶構造のデコードを始めとして、本研究課題は他の ACT-I 研究者の皆様との共同研究のもとで最新の機械学習技術を発展させた技術開発の上にも実現されたものであり、挑戦性が高い研究と考える。

5. 主な研究成果リスト

(1)論文(原著論文)発表

- | |
|---|
| 1. Yuta Suzuki, Hideitsu Hino, Takafumi Hawaii, Kotaro Saito, Masato Kotsugi, Kanta Ono. Symmetry prediction and knowledge discovery from X-ray diffraction patterns using an interpretable machine learning approach. Scientific Reports. 2020, 10, 21790. |
| 2. Yoshihiko Ozaki, Yuta Suzuki, Takafumi Hawaii, Kotaro Saito, Masaki Onishi, Kanta Ono. Automated crystal structure analysis based on blackbox optimization. npj Computational Materials. 2020, 6, 75. |
| 3. ほか投稿中 1 本、投稿準備中 1 本 |

(2)特許出願

研究期間累積件数:1 件(出願準備中)

(3)その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

学会発表

1. 鈴木雄太, 尾崎嘉彦, 羽合孝文, 齊藤耕太郎, 大西正輝, 小野寛太, "ブラックボックス最適化を用いたリートベルト解析の自動化", 放射光学会 学術講演会, 10, Jan. (2021)
2. Yuta Suzuki, Hideitsu Hino, Takafumi Hawaii, Kotaro Saito, Kanta Ono, "Machine learning approach for on-the-fly crystal system classification from powder x-ray diffraction pattern", TMS 2020 Annual Meeting, 23-27, Feb. (2020)

招待講演

1. 鈴木雄太, 奥野智也, 佐々木勇和(3 人による分担発表), "深層学習と物質探索", 応用物理学会第 1 回 インフォーマティクス応用研究グループ 研究会, 26, Nov. (2020)
2. 鈴木雄太, 小野寛太, "量子ビーム計測データ解析における機械学習の応用", 日本表面真空学会 学術講演会, 19, Nov. (2020)
3. Yuta Suzuki, "Machine-learning-aided data analysis in X-ray diffraction and absorption for high-throughput measurement", International Young Researchers Workshop on Synchrotron Radiation Science 2019, 3-4, Sep. (2019)

解説記事

1. 鈴木雄太, 尾崎嘉彦, 小野寛太, "解説:情報学的手法を用いた触媒開発 ブラックボックス最適化による結晶構造解析の自動化", 触媒. 63, 294-298, (2021)
2. 奥野智也, 佐々木勇和, 鈴木雄太, "深層学習を用いた新物質探索に関するサーベイ", 情報処理学会論文誌 データベース. 13 3, 1-10, (2020)

書籍

1. 鈴木雄太, 小野寛太, "マテリアルズ・インフォーマティクスのためのデータ構築技術と材料開発へのアプローチ" 11 章第 8 節 "機械学習によるスペクトルデータ解析とその事

例”(分担執筆), 技術情報協会. (2021)

受賞

1. 鈴木雄太, 学生発表賞, 日本放射光学会 (2021 年 1 月)

プレスリリース

1. X 線回折パターンからの対称性予測における知識発見-熟練者の勘・コツの定式化に成功-
<https://www.kek.jp/ja/press/20201211/>
2. 結晶構造解析の自動化 -ブラックボックス最適化により熟練者を上回る解析精度を達成-
<https://www.kek.jp/old/ja/newsroom/2020/06/05/1942/>