

ニューラルネットワークを用いた化学変化プロセスの表現

鐘川 凌 福田 龍樹
北九州工業高等専門学校生産デザイン工学科

1. はじめに

これまでの新材料開発やその研究は、研究者個人の経験や思い付きによって様々な材料の合成や実験を行い、その特性を調査することを繰り返し行うことで実現されてきた。しかし、新材料開発の目的は、より優れた特性を持つ材料を開発することであり、要求される性能をもとに必要となる物質を推測し合成が行われるべきであるため、従来の既存の物質からなる材料の特性を調べる手法とは本質が異なっている。また、従来の試行錯誤的に開発を進める手法では、複雑な組成構造の組み合わせから目的の構造を探索することは困難であり、膨大な時間と手間がかかる。そのため、現代の材料開発には高度化、高速化が求められ、その鍵として注目されているのがマテリアルズインフォマティクスである。

2. マテリアルズインフォマティクス

近年、材料科学とデータ科学とを融合したマテリアルズインフォマティクス(以下 MI)が注目されている。様々な材料を利用して実験を行い、理論計算を行う従来の手法に比べ、物質特性をコンピュータ上で高精度に計算した材料データベースや人工知能を用いるこの手法は時間やコストの削減、材料開発の効率化などが期待される[1]。

ところが、MIにも課題となる点があり、主な問題点として材料データが不足しているという点があげられる。MIを進めるためには、材料の特性を示すデータやその材料の製造過程のデータが大量に必要となるが、材料開発の過程で悪い結果を示したデータについては論文などで公開されず、企業や研究室で秘匿されているものが多い[2]。

そのため、少ないデータから効率的にデータを解析、学習できるツールを開発する必要がある。近年、データからその特徴を学習するディープラーニングの発展が著しいが、その中でGAN(Generative Adversarial Networks)と呼ばれる技術が注目されている。

3. GAN

機械学習の分野においてニューラルネットワークという概念がある。情報と情報のつながりに重みをつけ、その情報についての特徴を算出するこの概念は機械学習の基本となっている。この概念を用いたモデルにGAN[3]がある。データを生成する生成器とそのデータを本物かどうか識別する識別器の二つのニューラルネットワークでロス関数を共有させ、生成器はロス関数を小さく、識別器はその値を大きくすることを目的に互いに学習させるこのモデルは、

最終的に生成器がより本物に近いデータを生成することができるようになる。

図1にGANを用いて生成した画像について示す。これらの画像は、MNISTと呼ばれる手書き数字画像を集めたデータセットを使用し学習を行い、生成した画像である。図の各回数は、その画像を生成した時点での学習回数を示し、学習回数を増すごとにより本物に近く、優れた画像が生成できることが確かめられる。

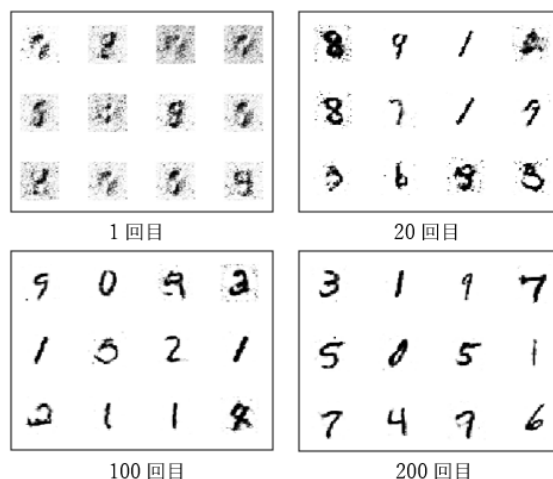


図1. GANを用いて生成した画像データ

4. 今後の展望

これまでで述べた通りGANはデータの特徴を算出し、学習を行うことで本物に近いデータを生成する能力を持つ。また、GANは二つのニューラルネットワークを競合させて学習を行うため、通常の機械学習の手法に比べて少ないデータ数での学習が可能となる。したがって、MIに必要なとされる材料データの不足の解決や効率的なデータ解析モデルの作成にGANは適しているといえる。

今後はGANを用いて材料データの解析、学習を行うモデルの作成を行い、MIに活用することを目指す。

参考文献

- [1] 伊藤 聡, “日本のマテリアルズインフォマティクス研究,” 人工知能学会誌, vol.34 no.3, pp.325-329, May 2019.
- [2] 畑中 美穂, “化学におけるマテリアルズインフォマティクスの現状と課題,” 人工知能学会誌 vol.34 no.3, pp.351-357, May 2019.
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, and Y. Bengio, “Generative adversarial nets,” Advances in Neural Information Processing Systems, pp.2672-2680, Dec 2014.