

人工知能が芸術的な色の表現で人間に勝つ日は来るのか

—色の表現を創生するディープラーニングを使用したインタラクティブアートに関する基礎研究—

研究代表者 芸術学部 インタラクティブメディア学科 教授 久原 泰雄

本稿では、私立大学研究ブランディング事業「色で明日を創る・未来を学ぶ・世界を繋ぐ KOUGEI カラーサイエンス&アート」における研究テーマとして採択された「色の表現を創生するディープラーニングを使用したインタラクティブアートに関する基礎研究」の進捗を報告するとともに、ニューラルネットワークにまつわる背景や芸術分野への応用の展望を示す。

1. 背景

人工知能(Artificial intelligence :AI)は、様々な分野において成果を上げている。グーグルのアルファ碁は自己対戦学習によって人間のプロ囲碁棋士に勝利し、IBM のワトソンは膨大な文献の学習によって専門分野の問題を解決することができる。自動運転車、ドローン宅配、接待ロボットなど、今後、ますます実用面での利用拡大が期待されている。

AI の中で特に注目に値するのはニューラルネットワークを利用したディープラーニングである。ニューラルネットワークは、理論をアルゴリズム化する代わりに、多くの経験を学習することによってヒューリスティックに直感的に問題解決を行う。したがって、ニューラルネットワークは、理性や論理以上に感性や直感が重視される芸術分野への応用が興味深いと言える。

2. ニューラルネットワーク概観

ニューラルネットワークの研究は決して目新しいものではない(表 1 参照)。1949 年に Hebb[1]は、神経細胞を繰り返し刺激するとシナプス結合の伝達効率が変化し学習する仕組みを定式化した。1957 年に Rosenblatt[2]がシナプス結合の伝達効率を調整して最適化を行うパーセプトロンを開発した。パーセプトロンは、最初のニューラルネットワークであり、その学習機能が脚光を浴びた。しかし、1969 年に Minsky[3]がパーセプトロンは線形分離問題のみ有効であり、非線形問題は学習できないことを示すと、ニューラルネットワークの研究は下火となった。

1980 年代に入り、ホップフィールドネットワーク、ボルツマンマシンなど有効なニューラルネットワークが開発されはじめた。1986 年の Rumelhart[4]が開発したバックプロパゲーションは複数の階層と最急降下法によるニューラルネットワークで、非線形問題を解決した。この手法によって、様々な分野でニューラルネットワークの応用が可能になった。例えば、1996 年に筆者ら[5]は、バックプロパゲーションを酵素機能と活性部位予測に応用した(図 1 参照)。とはいえ、大量の学習データを処理するには、当時の計算機の処理能力では不十分であり、学習に膨大な時間がかかるため、ニューラルネットワークの応用には限界があった。

表 1 ニューラルネットワークの変遷

年	出来事
1949	ヘッブの学習則
1958	パーセプトロン
第1次 冬の時代 (XOR 問題)	
1982	ホップフィールドネットワーク
1985	ボルツマンマシン
1986	バックプロパゲーション
第 2 次 冬の時代 (計算能力の限界)	
2016	アルファ碁

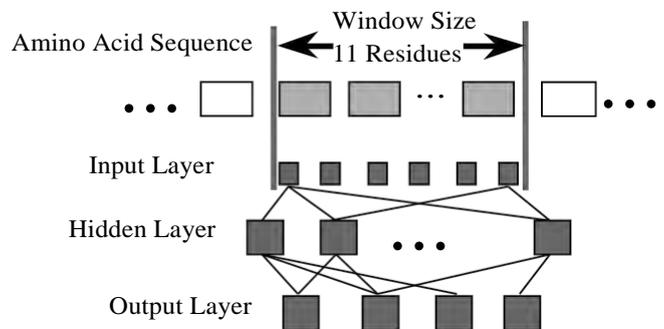


図 1 アミノ酸配列から酵素機能予測をするニューラルネットワークの応用例

2000年代に入り、コンピュータの性能が著しく進歩すると、ニューラルネットワークによる大規模な学習が徐々に現実味を帯びてきた。2010年代には高性能なGPUが開発され、計算能力が飛躍的に伸びると同時に、インターネットによるビッグデータが利用できるようになり、ディープラーニングと呼ばれる大規模な多層構造のバックプロパゲーションが構築され、驚異的な学習性能を発揮するようになった。2016年にはアルファ碁が人間の囲碁チャンピオンを負かすほどになった。アルファ碁は、戦法や定石などを論理的にプログラミングされていない。むしろ大量の対局データをニューラルネットワークによって学習し、論理よりも、直感に優れたAIとなっている。

3. 二人ゼロサム完全情報確定ゲームへの応用

囲碁のような二人ゼロサム完全情報確定ゲームは、理論的に完全な先読みが可能であり、双方の対戦者が最善手を打てば、必ず先手必勝、後手必勝、引き分けのいずれかが決まる。とはいえ、局面の変化が膨大になれば、現実的な時間内で完全な先読みを行うことができないためゲームとして成立する。ちなみに五目並べは先手必勝、4x4と6x6のオセロは後手必勝である。サイコロのような偶然性やトランプのような隠し札が存在せず、対戦者の実力のみで勝敗が決まるため、人間とコンピュータの能力を比較することに適している。表2に代表的な二人ゼロ和完全情報確定ゲームの局面数と人間に勝ったソフトウェアと時期を記した。現在では、いずれの場合もコンピュータは人間に勝っている。

表2 完全情報確定ゲームにおけるコンピュータの実績

ゲーム	局面数	コンピュータ
チェッカー(8x8)	10^{30}	2007 解決(引き分け)
オセロ(8x8)	10^{60}	1997 Logistello
チェス(8x8)	10^{120}	1997 Deep Blue
将棋(9x9)	10^{220}	2013 Ponanza
囲碁(19x19)	10^{360}	2016 Alpha GO

チェッカーは2007年にSchaefferらによって完全解析され、対戦者双方が最善を尽くした場合、引き分けになることが証明された。仮に1秒間に1億(10^8)の局面を解析できるプログラムを用いるとして、8x8のオセロの 10^{60}

の局面を解析すると、完全解析には以下の時間を要する。

$$10^{60} / 10^8 = 10^{52} \text{ 秒} \approx 3 \times 10^{44} \text{ 年}$$

宇宙の年齢が約138億(1.38×10^{10})年であることを考えると膨大な計算時間が必要であるかがわかる。チェッカー、オセロ、チェスはゲームが終盤になるほど局面の変化が減少するため、通常の木探索アルゴリズムと評価関数に基づく手法が有効である。一方、将棋は終盤に向かうと局面の変化が増えるため、木検索によって解を見つけることが困難になる。囲碁は盤面が $19^2=361$ と大きいことに加え、白と黒の配置パターンという2次元の画像から直感的に場面の優劣を判断することから、形勢判断の評価関数を作ることが難しい。そのため、ディープラーニングによるパターン学習が有効である。



図2 HPC システムズ GPGPU ワークステーション

4. 開発環境

ディープラーニングの開発環境として、Café, Chainer, TensorFlowなどが有名である。いずれもWindowsやMacOSで使用できるが、LinuxのディストリビューションであるUbuntuでの開発が効率的である。またニューラルネットワークの最適化は大量の行列計算を行うため、通

常のプロセッサに加えて NVIDIA の GPU が必須である。さらに大規模な学習を実行するには、通常のパソコンでは能力不足は否めず、専用の GPGPU (general-purpose computing on graphics processing units)を使用する必要がある。図 2 に本研究で使用した HPC システムズ GPGPU ワークステーション HPC5000 -XBWGPU4TS を示す。これには、Xenon プロセッサに NVIDIA Tesla GPU が 4 基搭載されている。日単位の学習時間が分単位で完了する性能を発揮する。

5. 芸術分野への応用

ディープラーニングは美術、音楽、文芸などの芸術分野にも応用されている。グラフィックスの分野では、コンテンツ画像をスタイル変換する手法が注目されている。

Gatys らの開発した A Neural Algorithm of Artistic Style[7]は、元となるコンテンツ画像と色彩や画風などの情報を持つスタイル画像の 2 種類の画像からコンテンツ画像に描かれた対象物の輪郭を保持しつつ、色彩や描画方法などの画風をスタイル画像から学習して変換する。松元らはこのアルゴリズムを chainer-gogh[8]として実装し、github にて公開している。

スタイル変換ニューラルネットワークは色彩や画風として用いるスタイル画像を学習させたら、ニューラルネットワークによって任意の画像を変換することができるため、様々な応用例が考えられる。本稿では、本学の元学長であり、イラストレーターでもある若尾真一郎氏による人物をモチーフとしたイラストレーションの画風を学習し、モノクロの人物画像を若尾イラスト風に変換するプログラムを作成した。若尾氏のイラストは独特の画風を持つが、決して一様ではなくバリエーションに富んでいる。今回はスタイル画像として画風の異なる 4 種のイラストを用いた。コンテンツ画像としてモノクロ人物写真を用いた。ディープラーニングのモデルとして、VGG_ILSVRC_16_layers を用いた[9]。また、lam 値(コンテンツ画像とスタイル画像の荷重比率)を様々な値で試し、綺麗な画像が生成されるよう調整した。結果の一部を末尾の付録に示す。パラメータを調整すれば、普通のモノクロ人物写真をスタイル画像が持つ色彩や画風を反映した人物画像に変換することができた。これらのスタイル変換をリアルタイムで実行するならば、様々なインタラクティブアート作品に応用することが考えられる。

6. 結論と展望

すべて論理的な思考で解決できる二人ゼロサム完全情報確定ゲームでは、もはや人間は AI に勝つことはできない。しかしながら、芸術分野においては、AI はいまだ発展途上である。論理的に勝敗が判断できるゲームの場合、優劣の評価が容易である。一方、芸術分野の場合、作品の優劣を客観的に評価することが難しい。すなわち、ニューラルネットワークにおける評価関数の設計が困難であるため、効率よく学習することができない。現在、研究制作されているのは、既存の作家の作品を学習させたニューラルネットワークを用いて、与えられたコンテンツの作風を変換したり、自動生成したりするものが多い。芸術分野において、AI が人間に勝ることは、今後の新たな開発を待たなければならないであろう。

今後のディープラーニングについて、量子コンピュータのような従来のノイマン型コンピュータと異なる超並列型アーキテクチャの実用化に期待したい。ノイマン型コンピュータは、巡回セールスマン問題の場合のように、変数が増える指指数関数的に計算時間が増大し、現実的な時間内に解を導き出すことができない。一方、量子の重ね合わせ状態やトンネル効果といった量子力学の現象を利用する量子コンピュータでは計算時間が大幅に改善される。ニューラルネットワークのような超並列計算向けの計算機アーキテクチャが実用化すれば、ディープラーニングの能力は飛躍的に向上すると考えられる。

参考文献

- [1] D. Hebb, "The Organization of Behavior", New York: Wiley & Sons, 1949.
- [2] F. Rosenblatt, "The Perceptron--a perceiving and recognizing automaton", Report 85-460-1, Cornell Aeronautical Laboratory 1957.
- [3] M. Minsky, S. Papert, "Perceptrons: an introduction to computational geometry", MIT Press, 1969.
- [4] D. E. Rumelhart, J. L. McClelland, "Parallel Distributed Processing: Explorations in the Microstructure of Cognition", Cambridge, MIT Press, 1986.
- [5] Y. Kuhara, K. Shimizu, J. Doi, "Enzyme Function Prediction Using Neural Networks of Variable Topology", International Conference on Engineering Applications of Neural Networks, pp127-130, 1997
- [6] J. Schaeffer et al, "Checkers Is Solved", Science, Vol. 317, Issue 5844, pp1518-1522, 2007.
- [7] Leon A. Gatys et al, "A Neural Algorithm of Artistic Style", arXiv:1508.06576, 2015.
- [8] E. Matsumoto, "chainer-gogh", <https://github.com/mattyachainer-gogh>, 2017.
- [9] K. Simonyan et al, "Very Deep Convolutional Networks for Large-Scale Image Recognition", arXiv:1409.1556, 2014

付録

コンテンツ画像



スタイル画像 1



スタイル画像 2



スタイル画像 3



スタイル画像 4



生成画像 1

生成画像 2

生成画像 3

生成画像 4

lam (コンテンツ画像荷重 / スタイル画像荷重) \doteq 0.3



lam \doteq 0.03



lam \doteq 0.01

