

(報道発表資料)

2023.01.10

日本電信電話株式会社
国立大学法人東京大学

物理深層学習のための新たな脳型学習アルゴリズムを開発 ～新アルゴリズムで光ニューラルネットワークによる高速な深層学習を実証～

日本電信電話株式会社(本社:東京都千代田区、代表取締役社長:島田 明、以下「NTT」と)国立大学法人 東京大学(所在地:東京都文京区、総長:藤井 輝夫、以下「東京大学」)大学院情報理工学系研究科 國吉・中嶋研究室 井上克馬助教及び中嶋浩平准教授(次世代知能科学研究センター(AI センター)兼務)らの研究グループは、脳の情報処理から得た着想を基に、深層ニューラルネットワーク(※1以下、ニューラルネットワークを NN と略す)ならびに物理系を計算過程に用いる物理 NN(※2)に適した新たな学習アルゴリズムを考案し、その有効性を確認しました。これを高速な機械学習器として期待されている光を用いた物理 NN に適用して、学習過程を含めて物理 NN 上で効率的に計算可能であることを世界ではじめて実証し、物理 NN としても世界最高性能を実現しました。本成果は人工知能(AI)向けコンピューティングの電力消費や演算時間の大幅な低減につながるものと期待されます。

本成果は、12月26日に発行された英国科学誌「Nature Communications」に掲載されました。

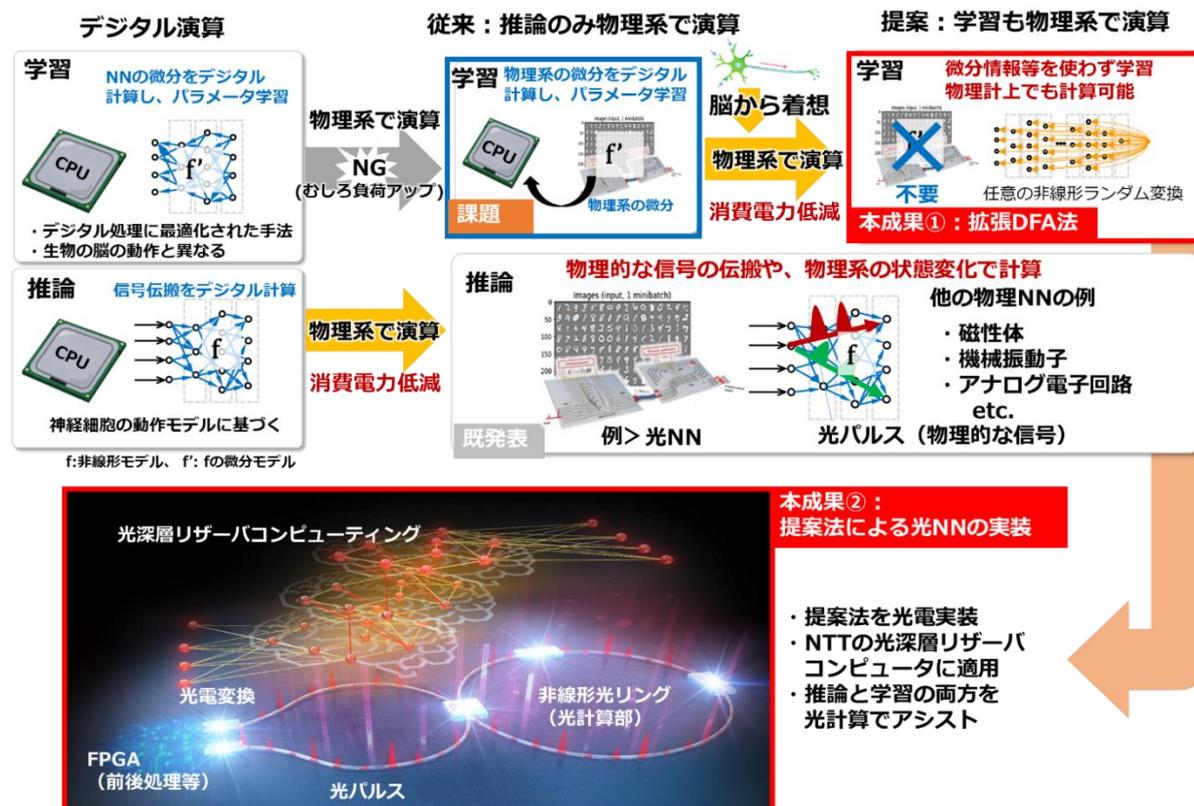


図1. (上)本成果の位置づけ。従来、高負荷な学習計算はデジタル演算で実行していましたが、本成果

の手法によって、学習部にも物理系による演算が組み込まれ、効率化が可能となりました。(下)実機実証用の光NNの概要。光パルスニューロンと、非線形光リングを回帰的な結合をもったNNと見立て、深層リザーバコンピューティング(※3)と呼ばれるリカレント NN(※4)を計算します。同一の光回路(※5)に出力信号を再度入力し、疑似的にネットワークの深層化を行っています。同一の系で推論用と学習用の計算を光演算でアシストしながら実行します。

1. 背景

機械翻訳、自動運転、ロボット制御などの高度な人工知能(AI)技術の基盤となる深層 NN では、演算量増大に伴い電力や計算時間が爆発的に増加しています。その伸びはデジタル計算機の性能向上の伸びを上回っており、深層 NN 技術を発展させるためには、計算機の性能の抜本的改善が喫緊の課題となっています。

様々な物理系の状態変化を計算過程として利用する物理 NN は、デジタル信号に依存しないニューロモルフィックデバイス(※6)の一種に分類され、爆発的に増大する深層 NN の演算に効率的に対処できる次世代の計算媒体として期待されています。特に光の伝搬・干渉による光演算を利用した NN(以下、光NN)は、低遅延性・低消費電力性・並列性を有し、そのため大幅な高速動作が理論的に可能で、既存のデジタル計算機の有力な代替候補としてニューロモルフィックデバイスの中でも特に注目されています。

深層NNの学習(※7)では、一般的に、誤差逆伝搬法(※8以下、BP 法)と呼ばれるアルゴリズムが用いられます。BP 法では深層NN中の学習パラメータ(※9)や非線形性、その微分値など、深層NNの情報の正確な把握が前提となっています。特に光 NN をはじめとする物理 NN に対して BP 法を適用するには、物理系上に構成された深層NNのすべての情報の把握、すなわち、高い精度での物理系の状態計測や物理シミュレーションによる微分応答の近似などが必要となり、より複雑な学習計算が要求されます。この複雑さのため BP 法の物理系上での実現は一般に困難とされ、通常デジタル計算機を介して学習計算が実行されますが、結果的に学習計算の大部分をデジタル演算に依存する形となり、物理 NN の拡張性ならびに有効性が大幅に制限されていました。加えてその制約の多さから、ある種の物理 NN とも捉えられる生物の脳の学習メカニズムとしても、BP 法の妥当性は懐疑視されており、生物の脳を含めた、より一般的に物理 NN 上で実行可能な学習アルゴリズムの開発が多くの研究者の関心の的になってきました。

2. 成果の概要

そこで本研究では、物理系の情報の正確な把握を必要としない、物理NNに適した「拡張 DFA 法」と呼ばれるアルゴリズムを新たに開発しました。

図 1(上)に本提案手法の概要と成果の位置づけを示します。研究チームは、脳の情報処理上で実現しやすい形で BP 法が改変されたダイレクト・フィードバック・アライメント法(※10以下、DFA 法)と呼ばれる学習アルゴリズムに着目し、物理 NN への実装に向けた拡張を行いました。提案法は、NNの最終層の出力と所望の出力信号間の差(誤差信号)にランダム要素を持つ行列で線形変換した数値を基に学習パラメータを更新して学習を進めます。この計算過程では BP 法では必要とされた物理系の状態計測や微分応答の物理シミュレーションによる近似は必要がありません。加えて、この計算は光回路をはじめとする物理系上で実行可能であり、推論だけでなく学習も物理系で効率的に計算できます。また研究では、この新しい学習手法が、物理実装で実現されているNNのモデ

ルだけでなく機械翻訳などで実際に使われている先端的な深層 NN モデルを含めた様々な機械学習モデルに対しても適用可能であることを明らかにしました。さらに原理実証として、推論と学習の計算を光演算によってアシストする光NNを構築(図1(下))し、従来困難であった光 NN の学習を光演算で効率的に実行可能であることを世界ではじめて示しました。また、構築した光NNを用いて、ほかの様々な物理現象を活用する物理 NN の中でも世界最高となる性能を実現しました。本成果により、学習から推論を通して光演算を組み込んだAIが実現され、今後顕在化すると予想されるAI向けコンピューティングの電力消費や演算時間の増大などの課題解決への貢献が期待されます。

3. 技術のポイント

(1) 物理 NN に適した「拡張DFA法」

脳の情報処理から着想を得て開発された DFA 法と呼ばれる手法に着目し、さらにこれを物理 NN での実装に適する形に拡張しました。図 2(a)に示すように、従来の DFA 法は、最終層の誤差にランダム行列を介して線形変換を行い深層 NN の学習を行います。BP 法と比較して物理 NN への適用性は高いものの、依然として物理 NN が動作する際に用いられる非線形変換およびその導関数の計算が必要でした。本提案手法ではこの部分を任意の非線形関数に置換できるように DFA 法を拡張しました。これにより BP 法で必要とされる物理系の状態計測や微分応答の物理シミュレーションによる近似、それに基づく逐次的な逆伝搬計算が完全に不要となります。学習過程の大幅な簡略化が実現され、物理 NN で事実上実現できていなかった学習を可能にしました。また、本手法の適用性を種々の深層学習モデルへ適用したところ、画像認識などで実際に使われている先端的なモデルから物理実装で盛んに研究が行われているモデルまで幅広い深層 NN モデルに対して適用可能であることが分かりました(図 2(b))。

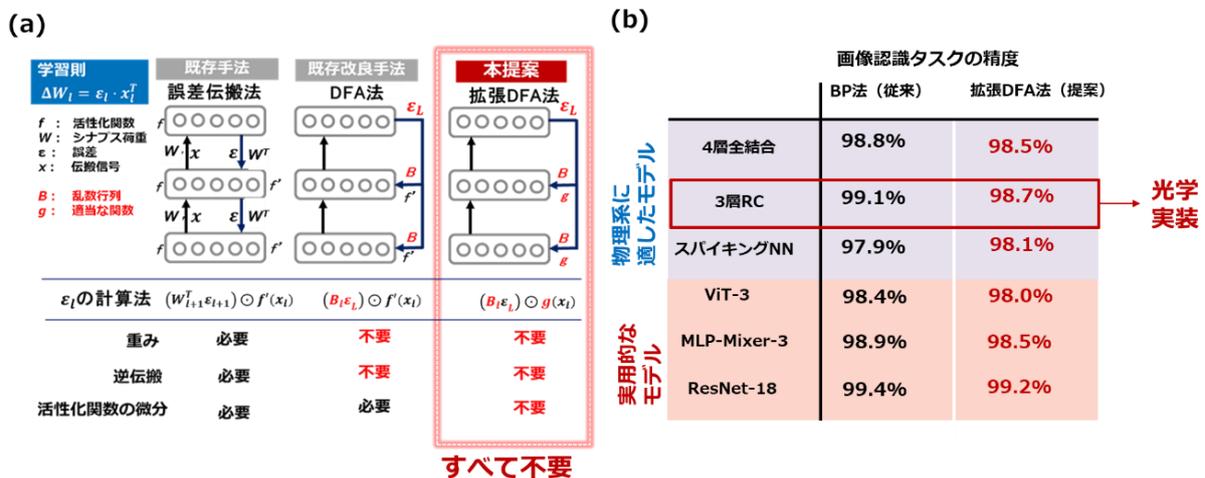


図 2:(a)既存学習法と提案学習法の概要、(b)種々のモデルでのベンチマーク結果。タスクには手書き数字の画像認識を用いている。提案手法の適用性検証には、物理実装で研究が盛んなモデル(青ハッチ部)や画像認識などで実際に利用されているモデル(赤ハッチ部)を用い、本ベンチマークに置いては、いずれでもBP法と同等の性能が確認できました。

(2) 光NN実機実験による動作実証

物理 NN への適用性を実機実験にて検証するために、深層リザーバコンピュータと呼ばれる深層 NNモデルを、光NN上に実装し、原理検証実験を行いました。学習部を含めた、光演算による計算加速の実機実証は世界で初めての試みとなります。この回路では、リザーバ層と呼ばれる回帰接続を持つ中間層の演算を、光パルスの強度をニューロンの応答と見立てて、非線形な光ファイバループで実行します。出力信号を再度同一の光回路に入力して、深層 NN モデルを計算できます。入力層や出力層の演算は電子回路上の演算で実行しますが、最も計算負荷の重いリザーバ層の計算を光学的に実行するため、計算規模が大きくなるにつれて計算時間や電力にスケールメリットが生じます。また、この回路のパラメータ設定を変更し、提案する拡張 DFA 法の計算の一部を光学的に処理できます。ハードウェアのみならず、上記の系を駆動するソフトウェアスタックを構築しており、ユーザは CPU や GPU などの一般的な計算機と同様に駆動できます。この光実装と提案アルゴリズムによって、従来非常に時間がかかっていた物理NNの学習過程が現実的な時間で実行できるようになり、本成果ではじめてCPUやGPUとも比較可能になりました。

構築した系を用いて、画像処理のベンチマークタスクの性能を比較した結果を図 3(a)に示します。提案手法並びに構築した光NNは、物理 NN としての最高性能を達成しました。図 3(b)に単位画像あたりの学習時間のニューロン数依存性を示します。比較として、CPUとGPUでの学習時間が示されています。小規模なネットワークモデルの場合は光ハードウェアへの情報転送の処理時間が律速するため、光演算による計算加速効果よりもそこまで大きな効果が得られていませんが、モデルの大規模化に伴って光NNの利用による顕著な学習加速効果が確認できました。これは、ニューロン数が大きくなるほど光 NN による演算が、従来型のデジタル電子演算に対して計算速度の観点で利点があることを示しています。また AI 計算の消費電力は効率と計算時間の積で決定されるため、演算の高速化は消費電力の低減に寄与します。このような学習部を含めた、光演算による計算加速の実機実証は世界で初めての成果です。今後は、効率の観点での優位性も合わせて検証を進めていきます。

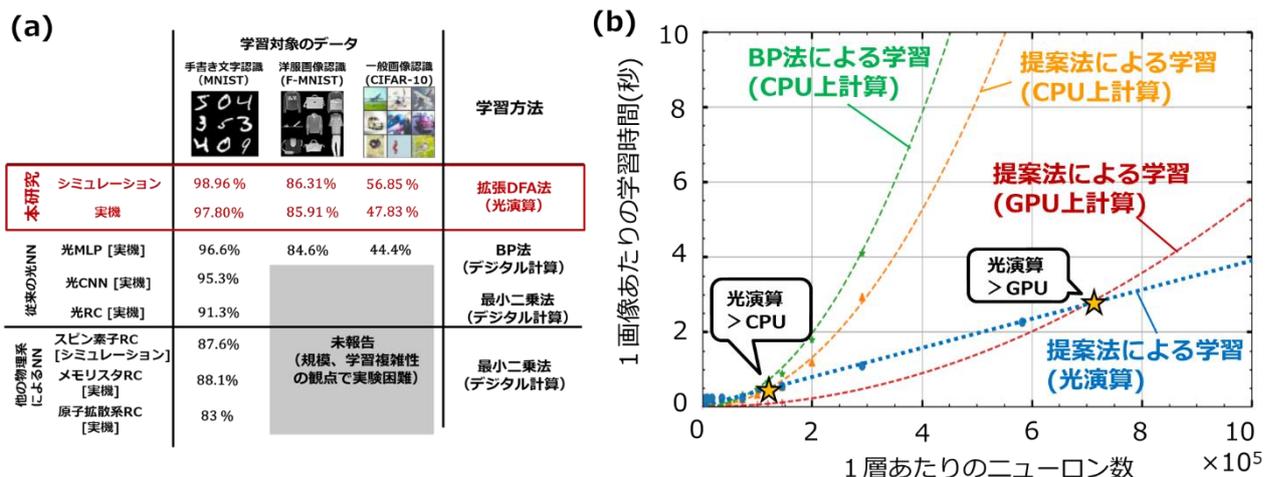


図 3:(a)構築した光 NN での画像認識タスク(手書き文字認識、洋服画像認識、一般画像認識)によるベンチマーク結果。参考として、光実装した畳み込み NN(CNN)、多層パーセプトロン(MLP)、リザーバコンピュータ(RC)と、他の物理系を利用した RC の従来研究を示します。従来は物理シミュレーションなどのデジタル計算に非常に時間のかかる BP 法か、十分な認識性能の期待できない最小二乗法を学習に利用していました。本研究で提案した拡張 DFA 法によって、物理系の情報を用いずに効率的かつ高性能

な学習が可能となりました。(b)単位画像当たりの計算時間のニューロン数依存性。光 NN の実測値(青線)のほかに、CPU(緑線、橙線)と GPU(赤線)での学習時間も比較として示します。ニューロン数の増加に伴って、光NNの利用による従来計算機(CPU,GPU)に対する学習加速効果が確認されました。

4. 今後の展望

本研究で提案した手法による、具体的な問題への適用性の検討も合わせて進めていくとともに、光ハードウェアの大規模小型集積を進めていきます。AI 技術の更なる発展は、ハードウェアの電力や計算時間が今後のボトルネックになることが懸念されています。これを抜本的に解決可能な光ネットワーク上での高速・低電力な光コンピューティング基盤を確立し、さまざまな社会課題の解決に適用することで安心安全かつ、ゆたかな社会の実現をめざします。

本研究への支援

本研究成果の一部は科学研究費助成事業(科研費)(JP18H05472, JP20J12815, JP22K21295)、科学技術振興機構(JST)戦略的創造研究推進事業 CREST(JPMJCR2014)、国立研究開発法人新エネルギー・産業技術総合開発機構(NEDO)の委託業務(JPNP16007)の結果得られたものです。

発表雑誌

雑誌名:「Nature Communications」(オンライン版:12月26日)

論文タイトル:Physical Deep Learning with Biologically Inspired Training Method: Gradient-Free Approach for Physical Hardware

著者:Mitsumasa Nakajima, Katsuma Inoue, Kenji Tanaka, Yasuo Kuniyoshi, Toshikazu Hashimoto, and Kohei Nakajima

用語の説明

※1 深層ニューラルネットワーク(深層 NN)

脳の情報処理から着想を得た数学モデルの一種。脳の神経網のような接続を、ネットワーク状に結合したノードと非線形変換で表現する。深層 NN は、多層の NN からなる数学モデルを指し、学習によるパラメータ最適化を通じて複雑な特徴を学習できるという特性を持つ。

※2 物理 NN

物理系を NN の計算素子として活用する情報処理フレームワークの一種。電流、磁気モーメント、構造変位、光強度など、自然界に存在する様々な物理量の物理系内での変化を NN の計算過程と捉えて情報処理を行う。物理系内の構造的な結合などがニューロン間の相互作用、非線形応答がニューロンの非線形性に相当する。

※3 リザーバコンピューティング

中間層に回帰接続を有するリカレント NN の一種。リザーバ層と呼ばれる中間層の接続をランダム

に固定することを特徴とする。深層リザーバコンピューティングでは、リザーバ層を多層に連結し、情報処理を行う。

※4 リカレント NN

ネットワークの中間層に回帰接続を有する NN。ネットワークの内部で過去の情報が巡回するので、時系列データなどの一連のデータの認識に優れる。画像データを処理する場合は、画像ピクセルを時系列状に並べて、学習や推論を行う。

※5 光回路

電子回路製造技術を用いて、シリコンや石英系の光導波路をシリコンウエハー上に集積した回路。通信では、光の干渉・波長合分波などにより光通信路の分岐や合流を行うが、光コンピューティングでは同様の原理で積和演算などの計算を光学的に行う。

※6 ニューロモルフィックデバイス

脳の情報処理を模倣した計算デバイス。元々は脳の構造とニューロンの発火メカニズムを物理デバイスで模倣したデバイスを指していたが、近年は深層 NN の計算を効率的に計算するデバイスも含まれる。本稿では後者の広義の意味で記載をしている。

※7 学習

深層NNに含まれる変数(学習パラメータ)に対して既知の入力に対する出力の組からなるデータの集まりを基に入力に対して正しい出力が得られるように最適化を行う。この最適化の過程を「学習」といい、最適化された学習パラメータを基に与えられた入力データから出力を計算する過程を「推論」と呼ぶ。

※8 誤差逆伝搬法(BP法)

深層学習で最もよく用いられる学習アルゴリズム。誤差信号を後方に伝搬しながら、ネットワーク内の重み(パラメータ)の勾配を求め、誤差が小さくなるように重みを更新する。逆伝搬過程に、ネットワークモデルの重み行列の転置操作や非線形性の微分演算が必要となるため、アナログ演算での上での実行は一般的に困難である。

※9 学習パラメータ

学習の対象となる深層NNに含まれる変数。

※10 ダイレクト・フィードバック・アライメント法(DFA法)

最終層の誤差信号を非線形なランダム変換することで、疑似的に各層の誤差信号を計算する手法。ネットワークモデルの微分情報などを必要とせず、並列的なランダム変換のみで計算できるため、アナログ演算と相性が良い。