

本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
 文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

シリコン光集積回路のみで作動するニューラルネットワーク演算技術を開発 — デジタル電子回路を補完する超低遅延、かつ低消費電力な光 AI 基本技術を確立 —

2022年6月29日

国立研究開発法人 産業技術総合研究所

日本電信電話株式会社

国立研究開発法人 科学技術振興機構(JST)

■ ポイント ■

- ・ 電子回路を使わず、シリコン光集積回路のみで作動する演算方式を考案し、その動作を確認
- ・ 光伝搬だけで演算できるため、超低遅延、かつ消費電力の少ない演算が可能
- ・ デジタル電子回路を補完する高速で低消費電力なAIアクセラレーターへの適用に目途

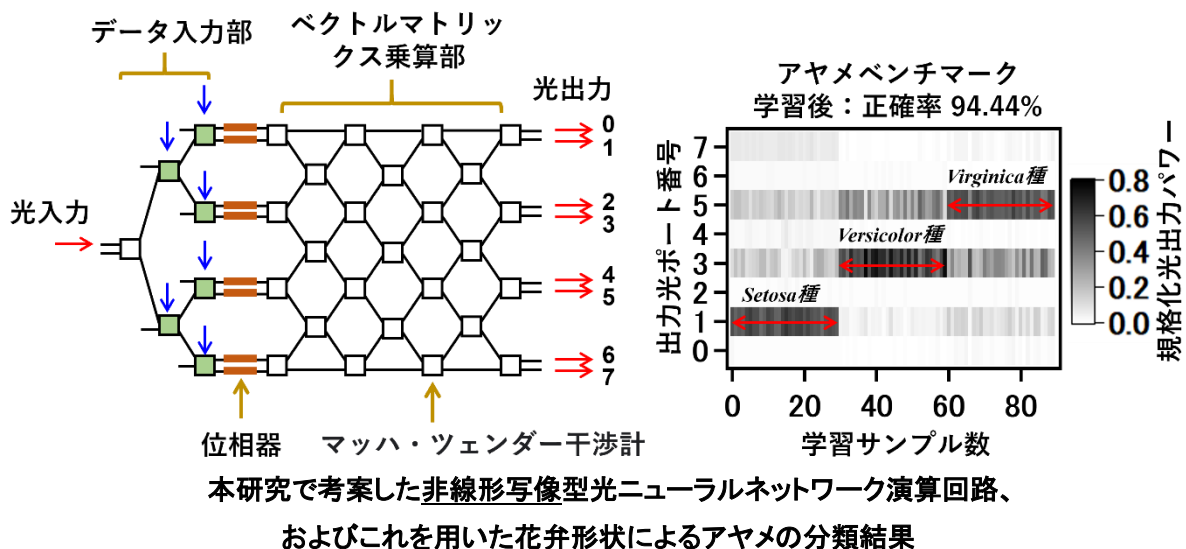
■ 概要 ■

国立研究開発法人産業技術総合研究所(以下「産総研」という)プラットフォームフォトニクス研究センターのコン グアンウェイ 主任研究員らは日本電信電話株式会社と共同で、国立研究開発法人科学技術振興機構の支援のもと、電子回路ではなく、シリコン光集積回路を使った超低遅延かつ消費電力の少ないニューラルネットワーク演算技術を開発した。

この技術は、光集積回路を用いて機械学習の演算を行う技術である。解析すべき多次元データの電気信号は光集積回路のそれぞれ異なる入力ポートに入力されて光信号に変換され、さらに光集積回路に組み込まれた多数の光干渉計を通過する際に演算が行われる。そして、複数の出力ポートの光強度分布として演算結果が出力される。

この技術を用いることにより、電気回路を経ることなく、光集積回路のみによるニューラルネットワーク演算が実現した。このニューラルネットワーク演算では、パラメーターが固定された光集積回路に光を伝搬させるだけで演算が完了するため、デジタル電子回路のような逐次スイッチングが不要となり、電子回路の千分の1以下の遅延時間、かつ数十分の1の消費電力での演算が可能となる。また、光回路では電子回路の10倍以上の高速なクロックが適用できるため、単位時間あたりのデータ処理量も大きくできる。これらの特徴により、この技術はデジタル電子回路を補完する AI アクセラレーターへの応用が期待される。なお、この技術の詳細は2022年6月30日に Springer Nature 社刊行の「Nature Communications」で発表される。

は【用語の説明】参照



本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

■ 研究の背景 ■

高度にデジタルトランスフォーメーション化された社会では、データセンターからエッジコンピューティング、自動運転、そしてコンシューマーデバイスまですべての情報機器において、幅広いAI処理の適用が必要とされている。AI処理システムの規模は毎年10倍程度の爆発的な拡大を示しており、既に1000億パラメーターの超大規模システムも登場している。このようなAI処理システムは、大量のデジタル演算プロセッサにより構築されているが、デジタル演算は演算規模の拡大により消費電力や演算遅延の増加が著しい。例えば、512個のGPUで構成されるAI処理システムは12万ワット以上の電力が必要である。また、エッジコンピューティングやロボット制御、自動運転などでは小規模なシステムが必要であるが、画像認識などのAI処理においてミリ秒程度の遅延を発生し、高速レスポンスが求められるこれらの応用にむけた課題となっている。

そこで近年、デジタル演算によらない、低消費電力で低遅延、そして大量のデータの処理が可能な高スループットAIアクセラレーターの研究開発が進んでいるが、その一つの候補として、光集積回路を用いた光ニューラルネットワーク演算が注目されている。光ニューラルネットワーク演算は、パラメーターが固定された光集積回路に光を伝搬させるだけで演算が完了する。デジタル演算のようなデバイスのスイッチングが不要なため、消費電力が少なく、また光集積回路チップを光が伝搬する時間で演算が完了するため、演算遅延は極めて小さい。しかしながら、現在の光ニューラルネットワーク演算回路は、光を用いた非線形性応答デバイスの集積が困難なため、光信号を電気信号に変換し、電子回路によるデジタル演算でこれを実現するハイブリッドな構成となっていた。したがって、現在の光ニューラルネットワーク演算は、消費電力および演算遅延の両観点において、そのメリットが活かされていない。そこで、光集積回路だけでニューラルネットワーク演算を実現する技術の開発が望まれていた。また、光演算回路の学習は、コンピューターによる事前学習が主流であるが、自らの経験に基づいた自律的な学習にむけては、回路実機に対する直接学習が必要であった。

■ 研究の経緯 ■

産総研では、次世代光集積回路技術であるシリコンフォトニクス技術の研究開発を進めてきた。シリコンフォトニクス技術は、シリコンエレクトロニクスの製造技術をベースとした光集積回路技術であり、微細加工性、集積性、経済性、そして省エネルギー性に優れ、近年の光集積回路の大規模化に必須の技術となっている。この技術を用いれば、光ニューラルネットワーク演算に必要な大規模光干渉計集積回路が実現可能である。さらに、シリコンフォトニクス技術により非線形性応答を担うデバイスを実現できれば、光集積回路のみでのニューラルネットワーク演算が可能となる。そこで、我々は光干渉計デバイスの駆動電圧に対する非線形性を用いた演算方式をシリコン光集積回路に実装して、光集積回路のみによるニューラルネットワーク演算を世界で初めて実証した。

本研究開発は、国立研究開発法人科学技術振興機構のCREST「新たな光機能や光物性の発現・利活用を基軸とする次世代フォトニクスの基盤技術」(JPMJCR15N4)および「情報担体を活用した集積デバイス・システム」(JPMJCR21C3)の支援を受けて行った。

■ 研究の内容 ■

本研究では、光干渉計デバイスの駆動電圧に対する非線形性を用いるために、非線形写像型のニューラルネットワーク演算方式を提案した。この演算方式では、データ入力部の光干渉計デバイスにより、解析すべきデータを高次元光複素振幅空間に非線形写像し、さらに多数の光干渉計から構成される光集積回路に光伝搬させることにより演算結果を得る。

本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
 文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

この演算方式を検証するための光集積回路をシリコンフォトニクス技術により製作した。図 1(a)に製作した回路チップ、図 1(b)に演算機能評価用に光ファイバーと電気配線を実装したモジュールの写真を示す。この回路は、シリコン導波路型マツハツェンダー光干渉計(MZI)および単体位相シフタを基本要素としたメッシュ構造となっている。MZI や単体位相シフタは導波路近傍に配置されたヒーターによる熱光学効果で動作する。この回路の動作原理について、分類演算を例に説明する。解析すべきデータは電気信号としてMZIに入力され、光信号に変換される。その際に、MZI の非線形応答により、入力データが高次元複素振幅空間に写像される。入力部を通過した光信号は、その後控えている積和演算部を通過し、そこでの演算を経て、分類の境界を与える高次元平面が算出される。最終的な分類演算結果は、最大光パワーを示す出力ポートの位置で示される。この演算原理のイメージを図 2 に示す。

デバイスにある干渉計パラメーターの設定は、機械学習により実施された。学習には細菌採餌最適化アルゴリズム(BFO)、あるいは前方伝搬アルゴリズム(FP)を利用し、回路実機に対し直接に学習を実施した。回路実機による直接学習は、外部コンピューターによる事前学習とは異なり、自らの経験に基づいた自律的な学習熟度の向上を可能とする技術であり、本研究の大きな成果の一つである。

この回路を用いた演算の例として、Iris flower classification と呼ばれる分類演算用ベンチマークの結果を図 3 に示す。この分類演算ベンチマークはアヤメの花弁サイズからアヤメの種類を判別するものである。この図の横軸はサンプル番号を示すが、サンプル番号 1~30、31~60、61~90 にはそれぞれ異なる 3 種類のアヤメ (Setosa 種、Versicolor 種、Virginica 種)を割り当てている。縦軸はこの回路の 8 個の出力ポートを示しており、1,3,5 ポートをそれぞれ 3 種のアヤメに割り当てている(ポート1は Setosa 種、ポート3は Versicolor 種、ポート5は Virginica 種)。分類結果は、最大光パワーを与えるポートが各アヤメ種に割り当てているポートと一致していれば正解となる(図3の右端の図参照)。学習前はサンプル番号と出力ポートの光パワーに相関はなく、分類はできていないが、90 サンプルの学習後は、約 94%の正答率で分類が可能となっている。さらに、学習に用いていない 60 サンプルに対しても分類を行い、約 97%の正答率も得られている。

分類演算の処理時間は、光集積回路を光が通過する時間であるが、この場合は 100 ピコ秒以下であり、デジタル電子回路演算の約千分の一であった。また回路パラメーターの設定に要したヒーター電力は約 360 ミリワットであり、同じくデジタル電子回路演算の数十分の一であった。さらに、データ入力干渉計を応答の遅いヒーター方式から数十 GHzで動作可能な PN 接合型高速シリコン光変調器に変更することにより、原理的には毎秒数 100 億回の高スループット演算も可能となる見込みである。

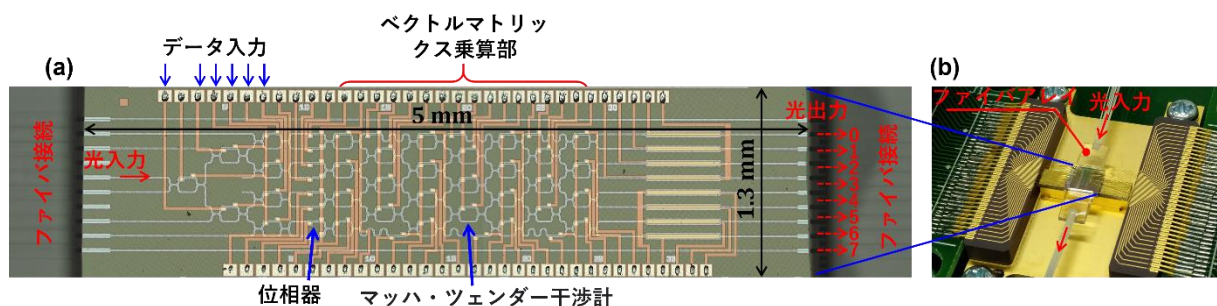


図 1 (a) 製作したニューラルネットワーク演算用シリコン光集積回路 (b) 評価用実装モジュール
 (図(a)は Nature Communications 誌に掲載された図面を和文に編集したもの)

本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
 文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

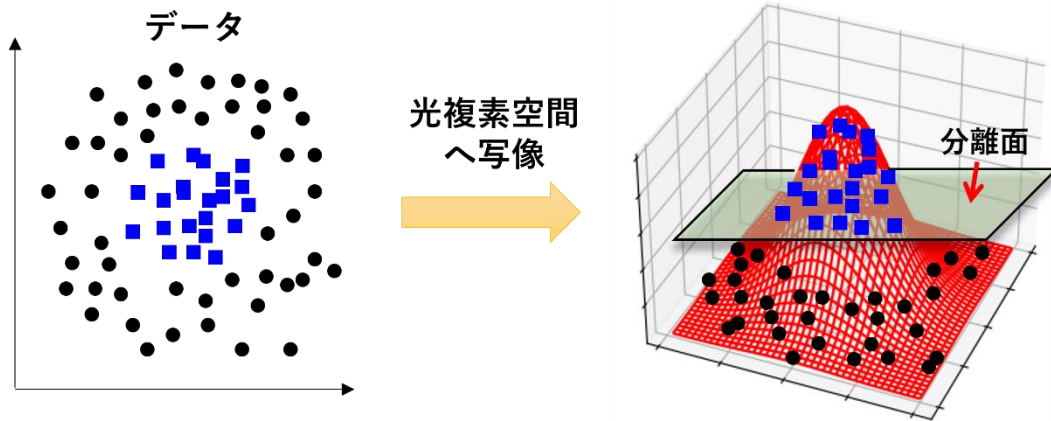


図2 非線形写像を用いた分類演算のイメージ図

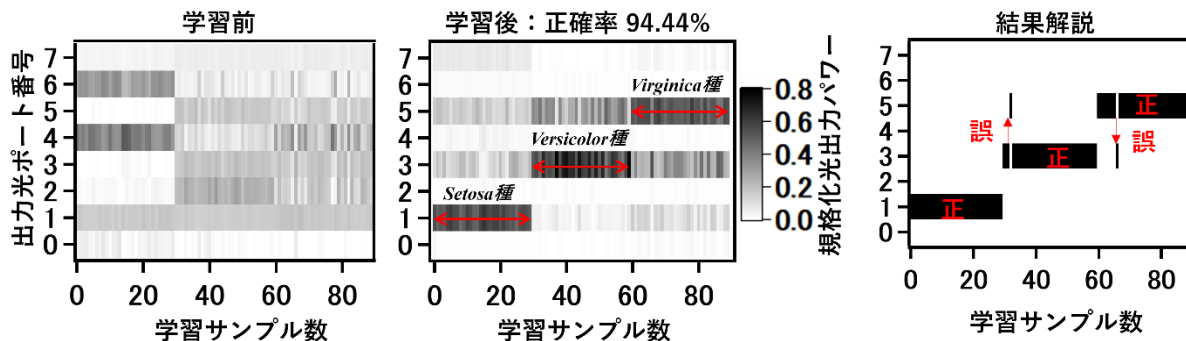


図3 アヤメの花弁形状による花種分類演算のベンチマーク結果。結果解説の図は学習後の結果を元に最大値を与えた出力ポートを示したものである。

(真ん中の図面は Nature Communications 誌に掲載された図面を和文に編集したもの)

■ 今後の予定 ■

今後は演算回路を大規模化し、より複雑な演算への適用性を確認していくとともに、高スループット化や学習機能の集積化にむけて、入力用高速光変調器や受光器の集積、デバイス駆動用、および学習制御用の電子回路の実装を進める予定である。また、今回提案した非線形写像型の演算方式をベースに、汎用データプロセッサや、再帰回路の付与による時間波形認識など、より幅広くかつ実用性の高い応用への適合性の確認を進める予定である。

■ 論文情報 ■

掲載誌: Nature Communications

論文タイトル: On-chip bacterial foraging training in silicon photonic circuits for projection-enabled nonlinear classification

著者: Guangwei Cong, Noritsugu Yamamoto, Takashi Inoue, Yuriko Maegami, Morifumi Ohno, Shota Kita, Shu Namiki, and Koji Yamada

本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

■ 本件問い合わせ先 ■

国立研究開発法人 産業技術総合研究所

プラットフォームフォトンクス研究センター

総括研究主幹 山田 浩治 〒305-8569 茨城県つくば市小野川 16-1 西事業所 TIA 棟
TEL:029-861-3699 FAX:029-861-5259
E-mail:yamada.koji@aist.go.jp

主任研究員 コン グアンウェイ 〒305-8569 茨城県つくば市小野川 16-1 西事業所 TIA 棟
TEL:029-861-5889 FAX:029-861-5259
E-mail:gw-cong@aist.go.jp

【機関情報】

国立研究開発法人 産業技術総合研究所
理事長 石村 和彦
URL:<https://www.aist.go.jp/>

[取材に関する窓口]

広報部 報道室

TEL:029-862-6216 E-mail:hodo-ml@aist.go.jp

日本電信電話株式会社

代表取締役社長 島田 明

URL:<https://group.ntt.jp/>

[取材に関する窓口]

先端技術総合研究所 広報担当

E-mail:nttrd-pr@ml.ntt.com

国立研究開発法人 科学技術振興機構 広報課

TEL:03-5214-8404 E-mail:jstkoho@jst.go.jp

【JST 事業に関する窓口】

国立研究開発法人 科学技術振興機構

戦略研究推進部 グリーンイノベーショングループ

嶋林 ゆう子

TEL:03-3512-3531 E-mail:crest@jst.go.jp

【用語の説明】

◆シリコン光集積回路

CMOS(相補型金属酸化膜半導体)電子回路製造技術を用いて、シリコンウエハー上に製作したシリコン光導波路をベースに構成した高密度集積光回路。

◆ニューラルネットワーク演算

神経系に存在するアーキテクチャーを模倣した電子回路や光回路などによる、ニューラルネットワーク等の人工知能アルゴリズムによる演算。

◆機械学習

人工知能処理の一種であり、学習データと呼ばれる経験データを使って学習し、データに潜むパターンや特徴を見つけ出し、その学習結果を用いて未知のデータに対して処理を行う。

◆AI アクセラレーター

人工知能(AI)処理の高速処理に特化した演算回路。

◆非線形写像

ある集合(定義域)に存在する要素を他の集合(値域)に写像する際に、各要素の線形演算では記述できない関数を用いて行われる写像。

◆エッジコンピューティング

ユーザーの近くに設置されたサーバー等を用い、クラウドにデータを送らずに、低遅延でデータ処理・分析を行う分散コンピューティングシステム。通信遅延が少なく、また計算負荷も分散されるため、リアルタイム性が高い。

◆光複素振幅

電磁波の電界状態を実部と虚部の二つの量で表現したもの。

◆光複素振幅空間

電界の実部と虚部の軸で張られた複素空間。複数の独立した電磁波が存在する場合、電磁波の数に応じた多次元空間となる。

◆マッハツェンダー干渉計

一つの光を二つに分岐し、それぞれの光の間に位相差を与えたのちに、再び合流させ光の干渉現象を得るデバイス。

◆熱光学効果

屈折率が温度によって変化する効果。

本プレス発表配布先【産総研】経済産業記者会、経済産業省ペンクラブ、中小企業庁ペンクラブ、資源記者クラブ、
文部科学記者会、科学記者会、筑波研究学園都市記者会【NTT】総務省記者クラブ

◆分類演算

統計学において、データを事物の共通な性質に基づく種類に分けるための演算。

◆細菌採餌最適化アルゴリズム(BFO)

餌場でより多くの栄養を得るために移動、再生等を行う細菌の採餌行動を数学的に定式化したアルゴリズム。

◆前方伝搬アルゴリズム(FP)

微小量の入力変化を与えた二回の順方向伝搬により勾配を計算し、ニューラルネットワークのパラメータを更新するアルゴリズム。