

どんな研究

光回路を用いたニューラルネットワーク（ONN）は超低消費電力と高速動作により注目されています。ONNは干渉計の多層構造で実現されるため、その数理モデルは干渉計表現行列の多段の積になります。このため、既存の自動微分による学習は多くの時間を要するという問題がありました。この展示では、ONNの高速学習法を紹介します。

どこが凄い

干渉計表現行列の多段積を細層構造ネットワークとしてモジュール化し、誤差逆伝播法で各細層の干渉計パラメータを更新するための専用偏微分式を用いる学習法を構築しました。回帰型ニューラルネットワークにおける文字認識実験で、本学習法は既存自動微分法に比べ、50倍以上の高速学習を達成しました。

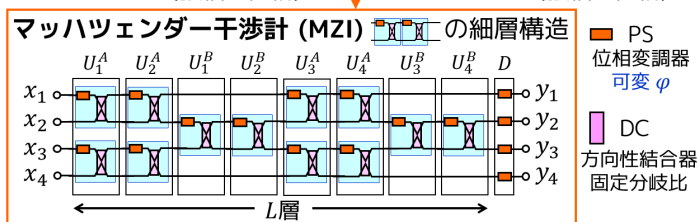
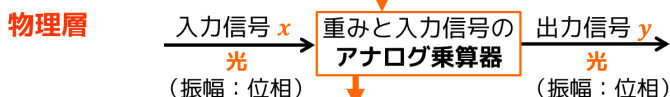
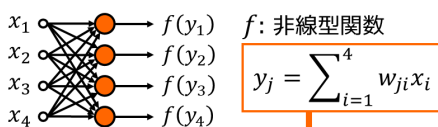
めざす未来

本学習法は、実光回路上のオンチップ学習や新しいデバイスへの応用という発展に加え、従来の膨大な電力を消費するAIではなく、超低消費電力で動作する光ニューラルネットワークによるグリーンITの実現に寄与します。

光ニューラルネットワーク (ONN) とは？

光干渉計アナログ回路で光信号と重み行列の乗算を行うNN

ニューラルネットワーク



数理モデル

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = D \dots \begin{pmatrix} U_1^B & & & \\ & U_2^A & & \\ & & U_3^A & \\ & & & U_4^A \end{pmatrix} \begin{pmatrix} U_{11}^A & & & \\ & U_{21}^A & & \\ & & U_{22}^A & \\ & & & U_{12}^A \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

U_p^R : ユニタリ行列 $R = A, B$ $1 \leq p \leq L/2$
 $\varphi_{p[q]}^R$: パラメータ $1 \leq q \leq dim/2$

$$U_{p[q]}^R = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\varphi_{p[q]}^R} & i \\ ie^{i\varphi_{p[q]}^R} & 1 \end{pmatrix}$$

学習における問題点

物理的要請：小規模光回路でデータ処理の必要有
 → 回帰型ニューラルネットワーク(RNN)の適用
 細層構造：各MZI層を線型回路と見なすことによる深層化

超深層ニューラルネットワークの学習

既存自動微分法 (AD) では
 計算グラフの大規模化により多大な計算時間が必要

高速学習法

ポイント1 専用偏微分式 (Customized derivatives: CD)

パラメータ更新式： $\varphi \leftarrow \varphi - \eta \left(\frac{\partial L}{\partial \varphi} \right)$ L : 損失関数 η : 学習率

1回の乗算で $(\partial L / \partial \varphi)$ を表すCDにより高速化

Forward $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\varphi} & i \\ ie^{i\varphi} & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Backward $\begin{pmatrix} \partial L / \partial x_1^* \\ \partial L / \partial x_2^* \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\varphi} & -ie^{-i\varphi} \\ -i & 1 \end{pmatrix} \begin{pmatrix} \partial L / \partial y_1^* \\ \partial L / \partial y_2^* \end{pmatrix}$

共役転置

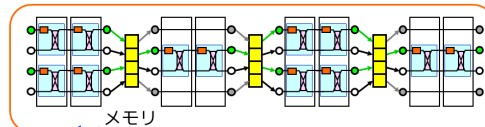
2つの値 $(x_1^*, \partial L / \partial x_1^*)$ の乗算で更新可能

$$\frac{\partial L}{\partial \varphi} = 2 \cdot \text{Im} \left(x_1^* \frac{\partial L}{\partial x_1^*} \right)$$

ポイント2 細層間高速アクセス (Pointer rewiring: PR)

細層間データ転送をコピー無しで実行することで高速化

多層構造をC++関数でモジュール化

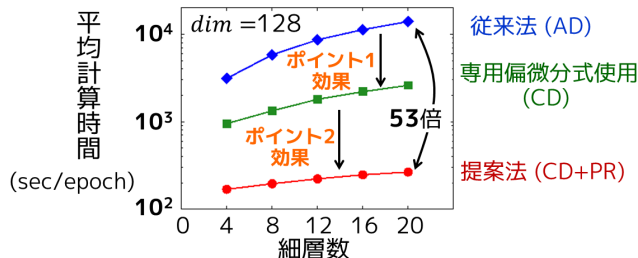


ポインター張替によるアドレスアクセスでデータ取得

効果

RNNの隠れ層を光回路で構成
 Pixel-by-pixel MNISTタスク(784時系列)

従来法と提案法の学習時間比較



関連文献

[1] K. Aoyama, H. Sawada, "Accelerated method for learning fine-layered optical neural networks," in Proc. of IEEE/ACM the 40th International Conference on Computer-Aided Design, 2021.

連絡先

青山 一生 (Kazuo Aoyama) 協創情報研究部 知能創発環境研究グループ
 Email: cs-openhouse-ml@hco.ntt.co.jp