ニューラルネットワークのシュレディンガー方程式による連続表現

Continuous Representation of Neural Network by Schrödinger Equation

NTT 先デ研 ○中島 光雅, 田仲 顕至, 橋本 俊和

NTT Device Technology Labs., Mitsumasa Nakajima, Kenji Tanaka, and Toshikazu Hashimoto

E-mail: mitsumasa.nakajima.wc@hco.ntt.co.jp

【背景】

近年、ディープニューラルネットワーク(DNN)を常微分方程式(ODE)としてみなす、Neural-ODE が機械学習分野で注目されている[1]。この枠組みでは、微分方程式 $x(t+1)=f[x(t),\theta,t]$を層の発展とみなし、出力演算や学習は ODE を解くことによって行われる[図 1(b)]。本稿では、Neural-ODE で取り扱う微分方程式を、光導波路中の伝搬を記述する波動方程式であるシュレディンガー方程式へと拡張する。学習 (設計) したネットワークは、物理量である屈折率を介して、物理的に実装可能である。

【ニューラルシュレディンガー方程式】

光導波路中の波動伝搬等をあらわすシュレディンガー方程式を Neural-ODE の枠組みで考える[図 1 (a)]。これをニューラルシュレディンガー方程式 (Neural-SE) と呼び、その処理手法を図 1(c)に示す。z 方向への波動伝搬が層の発展となる。この発展は、複素数型の畳み込みニューラルネットワーク(CNN)と等価であり、局所的な屈折率の分布が学習重みに相当する[2,3]。順伝搬はシュレディンガー方程式を解くことで計算できる。逆伝搬は、誤差逆伝搬法の連続近似と同等である Adjoint 法によって計算可能である[1,4]。いずれも有限差分法などの物理計算法を用いて計算可能である。このような系を考えることで、以下のような利点がある。

①局所的な物理量 (屈折率) を直接的に学習の対象として捉えられる。微小領域を"重み"として捉えられるため、ニューロモルフィック回路[5,6]の大規模集積化に有効である。

②トポロジ最適化を機械学習のフレームワークを利用して実行可能である。物理構造は、データ駆動で最適化される。

③ DNN と物理構造を同一のフレームワークで最適化できる。後段の DNN 処理に最適化された物理構造の探索(End-to-End 学習)など、新たな応用に結びつくと期待する。

【数値実験】

Neural-SE を汎用的な機械学習プラットフォームである Pytorch 上に実装した。計算手法には、数値安定性に優れるクランクニコルソン型の有限差分法を用いた。物理実装の容易性を鑑みて、吸収や利得は学習しないせず、局所的な屈折率の実部のみを学習している。手書き文字認識(MNIST)タスクで性能を評価したところ、99.3%と良好な精度を示した。当日は、End-to-End 学習や Pytorch 上での光回路のトポロジ最適化などの応用例も示す。

[1] T. Q. Chen *et al.*, "Neural ordinary differential equations," NeurIPS 2018, pp. 6571-6583 (2018).

[2] M. Nakajima *et al.*, "Neural Schrodinger Equation: Physical Law as Neural Network," *submitted.*

[3] L. Ruthotto *et al.*, "Deep neural networks motivated by partial differential equations," arXiv:1804.04272 (2018).

[4] M. B. Giles *et al.*, "An introduction to the adjoint approach to design. Flow," Turbul. Combust. 65, 393(2000).

[5] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits," Nat. Photonics 11, 441 (2017).

[6] M. Nakajima *et al.*, "Coherently Driven Ultrafast Complex-Valued Photonic Reservoir Computing," CLEO2018, SM1C.4.

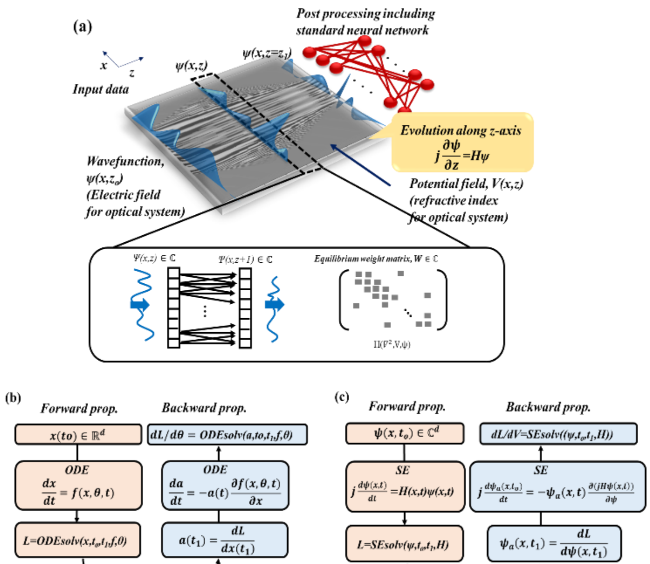

Fig. 1(a) Concept of neural network based on Schrodinger equation. (b) Diagram of neural ODE and (c) neural SE.