

令和 4 年 6 月 29 日現在

機関番号：31103

研究種目：研究活動スタート支援

研究期間：2020～2021

課題番号：20K23330

研究課題名（和文）高精度かつ安定な深層学習のための確率的擬等角活性化関数の構築

研究課題名（英文）Construction of Probabilistic Quasiconformal Activation Function for Highly Accurate and Stable Deep Learning

研究代表者

島内 宏和 (Shimauchi, Hirokazu)

八戸工業大学・工学部・講師

研究者番号：90759200

交付決定額（研究期間全体）：（直接経費） 1,300,000円

研究成果の概要（和文）：深層学習の普及が急速に進む中、その学習の高精度化、効率化、安定化等のために、活性化関数の構築に関する研究が進められている。本研究では、擬等角写像の技法と深層ニューラルネットワークに関する知見を融合することで、高精度かつ安定な深層学習に資する新しい活性化関数を構築することを目指した。研究成果として、確率的に変化するベルトラミ係数を持つ新しい活性化関数が得られた。構築した活性化関数は、複数のベンチマークデータセット上における基本的な深層ニューラルネットワークを用いた実験で、他のベースラインとした活性化関数と比較し高い性能を示した。

研究成果の学術的意義や社会的意義

深層学習による人工知能が、様々な分野において目覚ましい成果を挙げている。深層ニューラルネットワークの活性化関数は、構築されるモデルに関わる重要な要素の一つであり、その学習に大きな影響を与える。本研究で構築した確率的に変化するベルトラミ係数を持つ新しい活性化関数は、ベースラインとした既存の活性化関数と比較し、複数のベンチマークデータセット上で高い性能を示した。活性化関数のベルトラミ係数に摂動を与えるアプローチは、複雑かつ高い精度が求められる画像処理のタスク等における深層ニューラルネットワークモデルの構築に活用できる可能性がある。

研究成果の概要（英文）：The choice of activation functions significantly affects the performance of deep learning models. Therefore, the activation functions of neural networks have been widely researched. The aim of this research was to construct an activation function that contributes to highly accurate and stable deep learning. We focused on the Beltrami coefficient of the quasiconformal mapping: Beltrami coefficient represents the distortion of the mapping at each point. In this research, we constructed a complex-valued activation function that has a probabilistic Beltrami coefficient. The constructed activation function showed better performance compared with baseline activation functions in the experiment utilizing a simple deep neural network on several benchmark datasets.

研究分野：情報科学

キーワード：擬等角写像 ベルトラミ係数 ニューラルネットワーク 活性化関数 深層学習

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

データから規則を見出す機械学習の技術は著しく発展し、その手法の一つである深層学習による人工知能が、様々な分野において目覚ましい成果を挙げている。深層ニューラルネットワークの学習に際しては、層の種類や数、ユニット数、活性化関数、最適化法などを調整する必要がある。深層学習のハイパーパラメータの調整の自動化に関する技術の研究等も進んできているものの、複数の設定のパターンを試行する形となるため、試行の数だけ要する時間は長くなる。ニューラルネットワークにおいて、活性化関数は構築されるモデルに関わる重要な要素の一つであり、その学習に大きな影響を与える。学習の高精度化、効率化、安定化等のために、活性化関数の構築に関する研究が進められている。

層の深いネットワークを扱う場合には、決定論的な実数値関数である ReLU (Rectified Linear Unit)をはじめ、Leaky ReLU や Parametric ReLU などの ReLU の拡張、強化学習により探索された Swish (Sigmoid-weighted Linear Unit) 等が広く利用されている。また、決定論的な関数でない活性化関数としては、既存の決定論的な実数値関数に対し学習可能または学習不可能な重み付きの確率的に変化する摂動項を加える確率的活性化関数等が提案されている。また、複素数値のユニットおよび活性化関数を持つニューラルネットワークが、通信や音声、画像処理などの分野において有効な場合があることが知られており、近年その活用が広がっている。複素数値の活性化関数としては、ReLU を複素数値に拡張した Complex ReLU やほとんど至る所で正則かつ有界な活性化関数、振幅-位相型の活性化関数等が研究されていた。他方、擬等角写像という写像のクラスが、物理学におけるひも理論と関係の深いタイヒミュラー空間や複素解析関数による離散力学系(複素力学系)とともに研究されてきた。擬等角写像は、等角写像の一般化であり、可微分同相写像を含むより広いクラスである。これまで、擬等角性に着目した活性化関数の研究は見当たらない。

2. 研究の目的

本研究の目的は、擬等角写像の技法と深層ニューラルネットワークの活性化関数の先行研究により得られていた知見を融合することで、深層学習の高精化および安定化に資する新しい活性化関数を構築することである。

3. 研究の方法

Shridhar らの先行研究において、既存の決定論的な実数値活性化関数に対し、学習可能または学習不可能な重み付きの摂動項を加える形で定義される確率的な実数値活性化関数が構築されていた。これを複素化する、すなわち既存の決定論的な複素数値活性化関数に対し複素数値の確率的な摂動項を加えて構成する確率的活性化関数を考えると、これは既存の決定論的活性化関数に一種の線形等角写像を後から合成したものと見なすことができる。幾何学的には、活性化関数の出力に対し複素平面上でランダムな平行移動を加えていると捉えられる。

そこで、本研究では擬等角写像のベルトラミ係数に着目した。ベルトラミ係数は、定義域の各点における等角写像からの逸脱を表すものと解釈できる。活性化関数のベルトラミ係数に摂動を与える場合、それは幾何学的には複素平面上の各点において確率的な拡大・縮小・回転の操作を加えていると解釈できる。ベルトラミ係数は等角写像の後からの合成により不変であることを踏まえ、既存の決定論的複素数値活性化関数の前から確率的な擬等角写像を合成することで、ベルトラミ係数への確率的な摂動を与える方式を検討した。次の(1)、(2)、(3)のようなアプローチで研究を進めた。

- (1) はじめに、確率的に変化するベルトラミ係数を持つ擬等角写像を構成する。定義域におけるベルトラミ係数の絶対値により、対応する擬等角写像による像が退化する場合や反転する可能性があるため、その制限や摂動のスケーリングパラメータの導入等についても検討する。構築した擬等角写像に対し、既存の決定論的複素数値活性化関数を後から合成する形で定義される、確率的な複素数値活性化関数を構築する。
- (2) 1 で構築した確率的に変化するベルトラミ係数を持つ複素数値活性化関数の性能を実験的に評価する。既存の広く用いられている活性化関数および構築した活性化関数を用いて、ベースラインとするシンプルなアーキテクチャを持つ深層ニューラルネットワークを構築し、複数のベンチマークデータセット上でその性能を比較する。ここでは、全結合層によるニューラルネットワークおよび畳み込み層を用いたニューラルネットワークを採用する。なお、ニューラルネットワークの構築にあたっては、活性化関数の性能を評価するために、事前学習による初期値の学習や正則化を用いたモデルの制限、敵対的生成ネットワークによるデ

ータ拡張といった技法は排除する。実験で用いるデータセットとしては、画像識別および自然言語処理のタスクで利用される複数のベンチマークデータセットを採用する。また、データセットの訓練データが削減された場合の性能の変化についても検証する。

- (3) 2 で良好な実験結果が得られた活性化関数と、Shridhar らの先行研究の活性化関数の組み合わせについて検討する。先行研究の確率的な活性化関数は、複素化して考える場合には決定論的活性化関数に後から摂動項を持つ線形擬等角写像を合成していると考えられるが、2 で構築した活性化関数は決定論的活性化関数にベルトラミ係数に摂動を与えた擬等角写像を前から合成しているため、両者が与える摂動は本質的に異なるものである。そこで、構築した確率的に変化するベルトラミ係数を持つ活性化関数にさらに摂動項を加える、すなわち一種の確率的な線形擬等角写像を 2 で構築した活性化関数の後から合成する形で定義し、同様の形で実験的に評価する。さらに、基本的なニューラルネットワークに構築した活性化関数を用いた場合の近似能力や安定性について考察する。

4. 研究成果

はじめに、計算の観点からも扱いやすい、複素平面全体で指定された値をベルトラミ係数として持つ一種の線形擬等角写像を構築した。ベルトラミ係数に確率的なパラメータを導入し、それに対しネットワーク全体で共有される一つの学習可能なパラメータ、活性化関数ごとに別々の学習可能なパラメータ、学習不可能なパラメータという、3 種のスケールングのための重みをかける方式を検討した。この確率的な擬等角写像に、Complex ReLU などの既存の決定論的複素数値活性化関数を後から合成する形で、新しい活性化関数を構築した。

既存の広く用いられている決定論的活性化関数と今回構築した活性化関数を用いて、5 つの画像識別のベンチマークデータセット上で全結合層および畳み込み層を持つ深層ニューラルネットワークを構築し、その性能を比較した。実験では 3 回の独立試行のスコアの平均を用いた。本研究で構築した活性化関数は、他のベースラインとした活性化関数と比較し、複数のベンチマークデータセット上で高い性能を示した。また、自然言語処理のベンチマークデータセット上でも追加の実験を行い、良好な結果が得られている。

さらに、先行研究の確率的活性化関数の構成を踏まえ、構築した確率的に変化するベルトラミ係数を持つ活性化関数に対し、さらに確率的な摂動項を加える形の活性化関数を構築した。追加した摂動項に対しても、同様の 3 種のスケールングパラメータをかけた場合を検討した。上記の実験と同じベンチマークデータセットおよび深層ニューラルネットワークを用いた実験を行うとともに、ベルトラミ係数の摂動および追加された摂動項により豊かな特徴が抽出されうると考えたため、訓練データが削減された場合の性能も評価した。実験の結果、特に訓練データが少ない場合に、本研究で構築した活性化関数と先行研究の確率的活性化関数の組み合わせが高い性能を示した。

本研究で構築した活性化関数を用いた全結合ニューラルネットワークを考えた場合、学習可能な重みが零となるケースが含まれていることを考慮すれば、表現できる関数がなす関数空間は、構築した活性化関数の構成の中で用いる決定論的活性化関数を用いた場合の空間を含む、より広い空間をなすと考えられる。そのため、適切な決定論的活性化関数を構成の中で用いれば、良い近似を与える解が含まれると考えられるが、実験において見られた精度向上の理由の理論的な検討は今後の課題である。

なお、今回は構成の中で用いる決定論的な活性化関数として Complex ReLU を用いた場合を中心に検証したが、タスクによっては位相-振幅型の活性化関数等が有効な場合もある。タスクおよび活性化関数に応じ、それに特化したベルトラミ係数への摂動の与え方を検討するなど、さらなる展開の余地がある。

本研究の成果は、その過程で得られていた結果を論文としてまとめ、人工知能に関する国際会議 14th International Conference on Agents and Artificial Intelligence (Vienna, Austria, Online) の Deep Learning Session に採択され発表した。なお、同国際会議においては Neural Network Session の座長を務めた。また、本研究に関連してある種の正規化条件の下での擬等角写像の数値的構成手法や、擬等角写像を応用した一種の特徴抽出についても考察し、オンラインで開催された第 63 回 函数論シンポジウムおよび One-day Workshop on Applied and Computational Complex Analysis 2021 にて発表した。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Shimauchi Hirokazu	4. 巻 3
2. 論文標題 An Activation Function with Probabilistic Beltrami Coefficient for Deep Learning	5. 発行年 2022年
3. 雑誌名 Proceedings of the 14th International Conference on Agents and Artificial Intelligence	6. 最初と最後の頁 613-620
掲載論文のDOI（デジタルオブジェクト識別子） 10.5220/0010870500003116	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 2件/うち国際学会 0件）

1. 発表者名 島内宏和
2. 発表標題 区分線形写像とconformal weldingによる擬等角写像の近似
3. 学会等名 第63回 函数論シンポジウム（招待講演）
4. 発表年 2020年

1. 発表者名 島内 宏和, 須川 敏幸
2. 発表標題 等角写像の“指紋”と数値擬等角写像
3. 学会等名 One-day workshop on applied and computational complex analysis (ACCA2021)（招待講演）
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------