

早稲田大学大学院 先進理工学研究科

# 博士論文概要

深層ニューラルネットの  
積分表現理論

Integral Representation Theory  
of Deep Neural Networks

申請者

園田	翔
Sho	SONODA

電気・情報生命専攻 情報学習システム研究

2016年10月

深層ニューラルネットは、2012年頃から機械学習や人工知能の分野で急速に発展を続けている学習機械である。深層ニューラルネットの快挙は、大画像に対する一般物体認識タスクで人間と同程度のスコアを記録し、囲碁では「人類最強」とも呼ばれる棋士イ・セドル氏に勝利するなど、枚挙に暇がない。ニューラルネットは、神経細胞が繋がり合っただけで情報を処理する様子を抽象化した「脳の数理モデル」として、20世紀半ばに登場し、これまでに二度のブームを引き起こしている。深層ニューラルネットは第三次ブームの立役者である。

「深層」という修飾語は、中間層の数が従来のニューラルネットよりも多いことを強調している。ニューラルネットを深層化することで、内部の情報表現が階層化され、情報処理が効率化されることは、以前から予想されていた。しかし、古典的な学習法であるバックプロパゲーション (backpropagation) では、深層ニューラルネットを学習させることができなかった。原因は様々だが、例えば、層が深くなるに連れて、学習に必要な誤差信号が減衰し、学習が極端に遅くなるためである。深層ニューラルネットを学習させる技術を総称して、深層学習という。深層学習が立て続けに成功し始めたのは、2006年の Hinton や Bengio のプレトレーニングからである。

本研究では、深層ニューラルネットの中で何が起きているのか、なぜ深層にした方が良いのかという問題に対して、深層ニューラルネットの積分表現理論の開発を通じて問題解決を図る。深層ニューラルネットの内部では、タスクに有利な情報表現 (特徴量写像) が獲得されていると考えられている。情報表現を自動的に獲得するという意味で、深層学習は表現学習とも呼ばれる。しかし、深層学習はヒューリスティクスを多く含むので、実際に獲得される特徴量の素性は分からないことも多い。そもそも、浅いニューラルネットは任意の関数を近似できるほど表現力が高い (万能関数近似器) のに、なぜ深層にする必要があるのだろうか。

本研究が拠り所とする積分表現は、ニューラルネットの中間層素子に関する総和を積分に置き換えて得られる。これは中間層素子を積分核とする積分変換であり、双対リッジレット変換と呼ばれる。リッジレット変換は Radon 変換やウェーブレット変換との関係が深く、幾何学的性質や解析的性質がよく調べられている。通常のニューラルネットは、積分表現の離散化を通じて理解できる。積分表現理論は 90 年代に起きた第二次ブームにおいて、浅いニューラルネットの表現能力を調べる過程で成立した。残念ながら、深層ニューラルネットの積分表現理論は今日までほとんど調べられていない。中間層が二層以上ある場合には、単に積分核が入れ子になるだけで、中間層同士の関係をうまく定式化できないためである。

本研究の結果は二つに分けられる：浅いニューラルネットの積分表現理論と、深層ニューラルネットの積分表現理論である。浅いニューラルネットの理論では、ReLU と呼ばれる活性化関数に対応するように積分表現理論を拡張し、ニューラルネットと Radon 変換およびウェーブレット変換との関係を詳らかにし、さらに

積分表現を離散化してニューラルネットを学習する方法を提案した。

深層ニューラルネットの理論では、デノイジング・オートエンコーダー (denoising autoencoder; DAE) と呼ばれるクラスに対して、DAEを輸送写像とみなす方法で、積分表現を構成した。また、輸送写像の極限を調べることで、無限層ニューラルネットに相当する連続 DAE の性質を明らかにした。DAE はデータ分布のエントロピーを減らす方向に入力データを再配置する輸送作用があり、この作用は層を深くした方が顕著になることが分かった。従って、浅い DAE と深層 DAE とでは抽出される特徴量が異なることから、DAE においては積極的に深層化すべきであると言える。本研究の結果を深層学習のアルゴリズムに反映する方法の開発は、今後の重要な課題である。

本論文の構成は、第 1 章が本研究の概要と論文の構成の説明、第 2 章が関連研究と先行研究のサーベイ、第 3 章と第 4 章が本論を展開するうえでの準備、第 5 章から第 7 章が本論、第 8 章が本研究の総括である。第 2 章以降の各章の詳細は次の通りである。

第 2 章では関連研究および先行研究について、深層ニューラルネットと浅いニューラルネットの二つの観点で整理する。まず深層ニューラルネットについては、最初に全体の動向を概観する。次に、本研究の主題の一つである「深層ニューラルネットの中では何が起きているか」について言及している研究を整理する。本研究で取り扱う ReLU や DAE については独立に節を設けるほか、オートエンコーダーと対照的な表現学習の例として、畳み込みネットワークについても解説する。一方、浅いニューラルネットについては、まず 90 年代の結果を整理する。具体的には、万能関数近似能力を軸にして積分表現理論が登場するまでの経緯を説明する。続いて、積分表現理論以降に登場したリッジレット解析や学習理論について、その後の展開を整理する。

第 3 章では、本研究で用いる数学的な道具を整理する。具体的には、Fourier 変換や Radon 変換、ウェーブレット解析、拡散方程式、最適輸送理論の基本的な定理や公式を整理する。さらに、本論で展開される超関数や特異積分の計算について解説する。これらの計算には、これまでにまとまった解説が少なく、申請者が独自に計算した内容も含む。

第 4 章では積分表現理論について基本事項を説明する。本章は本論を展開するうえでの準備にあたるが、積分表現理論は本研究の要であり、申請者の考察も多く含むことから、独立に章を設けた。まず積分表現理論がリッジレット解析と等価であることを説明したあと、リッジレット変換が Radon 変換とウェーブレット変換の合成変換に分解できることを示す。これにより、リッジレット解析の幾何学的な意味付けが明らかとなる。最後に、リッジレット変換の離散化や、ベクトル値の場合の考え方を説明する。これにより、現実のニューラルネットと積分表

現との関係が明らかとなる。

第5章では浅いニューラルネットの積分表現理論を展開する。まず、深層学習において、ReLUと呼ばれる非有界な活性化関数が用いられる背景を簡単に説明する。これにより、深層ニューラルネットの積分表現理論を展開するためにはReLUを含む超関数によるリッジレット解析が必要であることが分かる。本章の前半では、超関数によるリッジレット変換が存在すること、および適当な条件の下で再構成公式（逆変換）が成り立つことを理論的に示す。後半では、リッジレット変換の具体例を解析的に計算し、さらに再構成公式の数値例を計算することで、理論の実効性を確認する。

第6章では深層ニューラルネットの積分表現理論を展開する。まず、DAEが登場した背景と、DAEの学習アルゴリズムを簡単に説明し、Alain and Bengioの変分計算によって学習アルゴリズムの停留点が陽に求まることを示す。続いて、得られたDAEが輸送写像とみなせることを説明する。本章の前半では、浅いDAEによる輸送の性質を調べる。後半では、三つの深層DAE（積層DAE、合成DAE、連続DAE）を導入し、深層DAEによる輸送現象を軸として深層DAEの積分表現理論を展開する。積層DAEは深層学習の一種であるプレトレーニングで現れる形式だが、解析が難しい。合成DAEは浅いDAEの合成写像であり、これ自体も輸送写像なので解析は比較的容易である。連続DAEは合成DAEの連続極限であり、無限層のニューラルネットに相当する。本章の主結果は二つある。まず、連続DAEによる輸送に伴って変形されたデータ分布（押出測度）が、逆向きの拡散方程式に従うことを示す。つまり、連続DAEはデータ分布のエントロピーを減らすようにデータ点を再配置する連続力学系である。次に、積層DAEと合成DAEの等価性を示す。つまり、積層DAEから得られる特徴量は、ある線形写像によって適当な合成DAEから得られた特徴量に変換できる。二つの主結果の系として、合成DAEと積層DAEはいずれも、層を重ねるに連れて連続DAEと類似の振舞いをするようになることが分かる。最後に、深層DAEの積分表現は、層毎の積分表現を合成したものとして得る。

第7章では積分表現を離散化することでニューラルネットを学習させる方法を説明する。再構成公式を離散化することで学習済ニューラルネットが得られる。離散化は離散フーリエ変換のように規則的な格子に沿って行うこともできるが、本章ではサンプリングによる方法を提案する。これは、データからリッジレット変換を推定し、得られた変換を確率分布とみなして、パラメータをサンプリングする方法である。リッジレット変換をパラメータ空間上の確率分布とみなしたものをオラクル分布と呼ぶ。人工データおよび実データに対してアルゴリズムを適用し、バックプロパゲーションに依らない学習が行えることを確認した。

第8章では本研究を総括し、今後の展望について述べる。

なお付録には、本文で省略した証明と、本研究の背景となる知識をまとめた。

## 早稲田大学 博士（工学） 学位申請 研究業績書

氏名 園田 翔 印

(2017年2月1現在)

種 類 別	題名、 発表・発行掲載誌名、 発表・発行年月、 連名者（申請者含む）
論 文 ○	<u>Sho Sonoda</u> , Noboru Murata, “Neural Network with Unbounded Activation Functions is Universal Approximator”, Applied and Computational Harmonic Analysis, Elsevier, (2015). (掲載決定)
講 演 国際会議 査 読 有 ○	<u>Sho Sonoda</u> , Noboru Murata, “Sampling Hidden Parameters from Oracle Distribution”, The 24th International Conference on Artificial Neural Networks (ICANN2014), Hamburg, Germany, September 15-19, 2014. (査読付論文)
講 演 国内会議	<u>園田翔</u> , “ニューラルネットの積分表現理論”, 第2回産総研人工知能セミナー「機械学習の理論的側面」, 台場, 2015年11月.  <u>園田翔</u> , 村田昇, “ReLU ネットワークの積分表現理論”, 2015年度 科学研究費シンポジウム「大規模複雑データの理論と方法論：最前線の動向」, 筑波, 2015年11月.  <u>園田翔</u> , 村田昇, “深層学習のリッジレット解析にむけた取組み”, 2015 RIMS 共同研究「ウェーブレット解析と信号処理」, 京都, 2015年11月.  <u>Sho Sonoda</u> , Noboru Murata, “Ridgelet Analysis of ReLU Network”, The 29th Machine Learning Summer School (MLSS2015), Kyoto, Japan, August 23 - September 4, 2015.
そ の 他	<u>園田翔</u> , 村田昇, “オラクル分布を用いたサンプリング学習アルゴリズム”, 第18回IBISML研究会, 筑波, 2014年9月.  (論文) <u>Sho Sonoda</u> , Noboru Murata, Hideitsu Hino, Hiroshi Kitada, Manabu Kano, “A Statistical Model for Predicting the Liquid Steel Temperature in Ladle and Tundish by Bootstrap Filter”, ISIJ International, 52(6):1086-1091, (2012).  (講演) 嶋田達之介, 松原拓央, <u>園田翔</u> , 村田昇, パトリシアオータル, 加藤真平, “LiDAR深度データを用いたCNNブレイクシーン認識”, 第18回情報論的学習理論ワークショップ(IBIS2015), 筑波, 2015年11月.  (講演) 金田有紀, <u>園田翔</u> , 日野英逸, 村田昇, “複数粒子フィルタとモデル選択を用いたEEGデータの電流ダイポール推定”, 第17回IBISML研究会, 沖縄, 2014年6月.  (講演) <u>園田翔</u> , 村田昇, 日野英逸, 進藤史裕, 北田宏, 加納学, “ブートストラップフィルタによる溶鋼温度分布の予測と制御”, 日本鉄鋼協会 第162回 秋季講演大会, 大阪, 2011年9月.  (プレプリント) <u>Sho Sonoda</u> , Noboru Murata, “Decoding Stacked Denoising Autoencoders”, Journal of Machine Learning Research, arXiv:1605.02832