

# 海馬の構造に着目した Restricted Boltzmann Machine の拡張モデル

## Extended Models of Restricted Boltzmann Machine Focused on Structure of Hippocampus

大澤 正彦<sup>1</sup> 今井 倫太<sup>1</sup> 山川 宏<sup>2</sup>

Masahiko Osawa<sup>1</sup>, Michita Imai<sup>1</sup>, and Hiroshi Yamakawa<sup>2</sup>

<sup>1</sup>慶應義塾大学 大学院 理工学研究科

<sup>1</sup>Graduate School of Science and Technology, Keio University

<sup>2</sup>株式会社ドワンゴ ドワンゴ人工知能研究所

<sup>2</sup>Dwango AI Laboratory, Dwango Co., Ltd.

**Abstract:** This research explores the information processing functions arising from the unique structure of the hippocampus. Specifically, two structural features are elaborated in this paper. The first is neurogenesis which takes place in the dentate gyrus and the other is the recursive neuronal projection seen in CA3. Here we refer to the representations of Restricted Boltzmann Machines (RBMs), one of the important forms of Deep Learning, and their constitutively extended models. In addition to the functionalities of ordinary RBMs, extended RBMs acquire novel features which may provide insight to the explanation of cognitive functions in the hippocampus.

### 1. はじめに

神経科学実験の成果として発見された知見は、人工知能関連技術の開発において多くの示唆を与えうると考えられてきた。一方で、機械学習をはじめとした様々なコンピュータサイエンス研究によって得られた知見も、神経科学において得られた知見に解釈を与えうる。

本稿では、海馬と呼ばれる脳領域の特徴的な構造が果たす情報処理機能を、さまざまな人工ニューラルネットワークモデルを用いた実験結果から考察する。このとき、近年注目を集めている深層学習の構成要素の 1 つとして知られている、Restricted Boltzmann Machine (RBM) の構造的拡張モデルが獲得する情報表現に主に着目する。

RBM がもつ特徴抽出器、連想記憶器、生成モデルとしての機能は、海馬の代表的な機能の 1 部としてあげられる場合がある。従って RBM は海馬の機能を考察する上で有望な可能性がある。

### 2. 海馬の構造

本章では、本稿で扱う海馬の知見をまとめる。

海馬の神経接続[1]を図 1 に示す。本稿では、図 1 に示したの中でも特徴的な構造をもつ歯状回と CA3 に注目する。

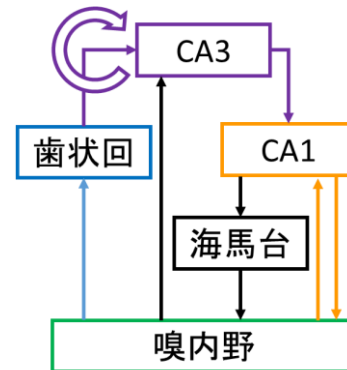


図 1 海馬の神経接続[1]

歯状回は主に嗅内野から入力を受け、CA3 へと情報を伝達する。成体におけるほとんどの脳領域で神経細胞は新たに発現しないことが知られているが、歯状回は例外的に多くのニューロンが新生していることが発見された[2]。新生ニューロンは可塑性が高い[3]ほか、パタン分離[4]や記憶の消去[5]など周辺のニューロンとは異なる機能性を持っているという知見が多く示されている。

CA3 は歯状回や嗅内野から受けた情報を CA1 へと伝達する。脳内でも特に強い再帰的神経投射が存在していることが知られている領域である[6]。

### 3. 海馬の構造に着目した

#### RBM の拡張モデル

本章では、2種類の構造的な特徴をもつRBMの拡張モデルと、それらを組み合わせた複合学習モデルを主に扱う。モデルが獲得する表現について考察することで、各構造が果たしうる情報処理に関する示唆を与えることを目指す。

#### 3.1 素子数が動的に増加する

##### RBM の拡張モデル

素子数が動的に増加するRBMの拡張モデルには、以下のモデルがある。

- ordered RBM[7]
- infinite RBM [7]
- incremental learning RBM (IL-RBM) [8]

RBM 以外では、以下のモデルは上記のモデルと共通する有用な示唆を与える。

- Resource Allocating Networks (RAN) [9]
- denoising Autoencoder を用いた逐次学習方[10]

#### 3.1.1 既学習情報を失わない逐次学習

多くの人工ニューラルネットワークの学習には、破壊的干渉と呼ばれる問題がある[11]。これは後から学習した情報が先に学習した情報を上書きする現象を意味する。従って人工ニューラルネットワークを用いて逐次学習を行う際には、注意が必要である。

海馬によって行われる記憶は、神経の可塑的变化を伴うが、逐次的に学習することが求められる。従って人工ニューラルネットワークと同様に破壊的干渉に準じた問題が起こる可能性が想定される。

新たな素子を動的に加えながら学習する方法は、破壊的干渉の対策として有効な場合があることが複数の文献で報告されている [9-11]。

したがって歯状回の新生ニューロンは、海馬での逐次学習で既学習情報の消失を防ぐ役割を担うことも可能性の1つとして考えられる。

#### 3.1.2 高速な学習

素子を動的に追加する学習手法の1部は、設定するハイパーパラメータや学習方法によって学習にかかる実行時間を大幅に短縮できる場合がある[15]。一方で、高速な学習によって得られた特徴表現は、低速な学習によって得られたものよりも汎化性能の点で劣っている場合が多い。

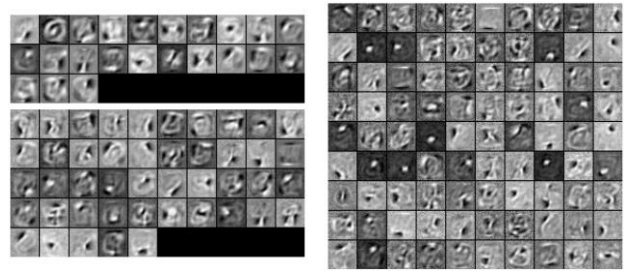
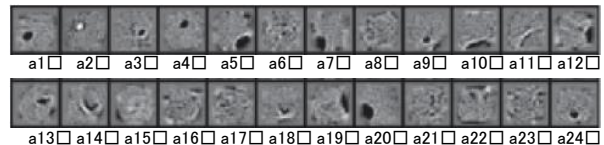
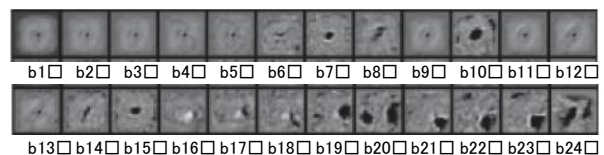


図 2 MNIST を学習した RBM の結合荷重の可視化。素子数: 23(左上), 45(左下), 90(右)



(a) RBM



(b) iRBM

図 3 2 値化された MNIST を学習した RBM (a) と infinite RBM (b) が獲得する表現の比較。(文献[8]で示された結果を一部抜粋・修正。)

海馬は頻繁に新皮質と対比して議論される。一般に新皮質は高い特徴抽出能力を持つが、記憶の定着には長い時間がかかる傾向がある。一方海馬は高速に記憶が定着する代わりに、特徴抽出の性能は新皮質ほど高くはないと考えられている。

新皮質と対応付けた海馬の知見と比較すると、歯状回の新生ニューロンは、海馬が実現している記憶の高速性に貢献している可能性が想像できる。

#### 3.1.3 抽象度の異なる表現の獲得

はじめに、素子数が異なる RBM が獲得する表現について簡単に説明する。図 2 に素子数の異なる RBM が獲得する表現の比較を示す。図 2 では、それぞれ 23 個、45 個、90 個の隠れ素子数をもつ RBM に対して、手書き文字データセットとして有名な MNIST を学習させた例である。

素子数が少ない場合には抽象度が高く、多くのデータに共通して現れる表現が獲得されている傾向がある。一方で素子数が多い場合には抽象度が低く、少数のデータに共通して現れる表現を獲得していると解釈できる。

素子数が動的に増加するモデルは、学習の段階によって獲得する表現の抽象度が異なる傾向がある。図3にRBMとinfiniteRBMが獲得する表現を示す。

図3(a)に示したRBMが獲得した表現には、抽象度に関する素子間での大きな差は見受けられない。一方図3(b)に示したinfiniteRBMは、抽象度が異なる表現が獲得されている。例えば、図中のb1に示した表現は、ほぼすべての学習データに共通して現れる表現と考えられる。一方b24に示した表現は主に1,2,7に対応するデータには現れやすく、その他のラベルに対応するデータには現れにくいことが想定される。モデルの性質上、例にあげたb1は学習初期段階で追加された素子であり、b24はb1よりも後に追加された素子と考えられる。

ただし、抽象度の異なる表現を獲得する学習器は、多層化した際の性能向上が限定的である場合が多い。特に深層学習においてRBMやAutoencoderなどの教師なし学習のモデルを段階的に積み重ねていく場合には、抽象度の低い表現のみを獲得するモデルによって段階的に抽象化することで、強い非線形性と汎用性をもつ表現を獲得できると考えられている。

### 3.2 再帰構造を有するRBMの拡張モデル

再帰構造を有するRBMには以下のものがある。

- Temporal RBM (TRBM) [12]
- Recurrent TRBM (RTRBM) [13]
- RNN-RBM [14]
- RBM-ESN [15]
- RBMを用いたESNの学習法 [16]
- Half RBM (HRBM) [17]
- Stacked HRBM (SHRBM) [17]

#### 3.2.1 時系列データの記憶・想起・特徴抽出

再帰的な神経投射が果たす情報処理の役割については多くの研究成果によって示唆が与えられてきた。例えば、再帰構造は連想記憶器にとって重要な役割を果たす。したがって古くから海馬は連想記憶器として対応付けられることが多い[18]。また、一般に時系列データの学習においても再帰構造は重要である。

文献[13-18]はいずれも時系列に依存したデータセットの学習や想起に関する課題を学習できることを示しており、RBMの学習は時系列の記憶や想起に応用可能と考えられる。

TRBM[12]やRTRBM[13]、RBM-ESN[15]は抽象的な2値画像からなる動画の記憶および想起ができる。



図4 実画像を用いた動画データセット

RNN-RBM[14]は音声データにも応用されており、生成モデルの性質を利用した簡単な作曲が可能となる例が示されている。

文献[16]では、事前に提案手法を用いた教師なし学習を行うことによって、少数の隠れ素子で性能が向上する例を複数の識別タスクを用いて示されている。さらに学習した時系列を逆順に想起することも可能である。

#### 3.2.2 作業記憶

文献[17]では、マウスの作業記憶機能を調べる際によく用いられる8方向放射状迷路課題を強化学習課題としてモデル化して評価実験に用いている。

HRBMやSHRBMを環境の特徴抽出器として利用すると従来手法よりも性能が向上した。従って、教師なし作業記憶機能を有する可能性が示唆されている。

### 3.3 海馬の構造に着目した複合学習器

本節では素子数が動的に増加するモデルと、再帰構造を持つモデルとを組み合わせたモデルの特徴について説明する。

はじめにIL-RBMに時系列データを入力した場合の性質について説明する。IL-RBMはある素子が追加されたタイミングが早いほど多くのデータに共通した特徴量を抽出する性質を持つ。このとき時系列データを学習すると、初期に素子は時間変化が小さい特徴量をよく学習する。一方学習後期に追加された素子は、時間変化が大きい特徴量を抽出する[15]。つまりSlow Feature Analysisのように、時間変化の大きさに応じた特徴抽出とも解釈できる。

3.1.3項ではIL-RBMなどの動的に素子数が増加するモデルの上に従来の深層学習に用いるRBMやAutoencoderなどの教師なし学習モデルを積み重ねてもほとんどの場合は精度の向上が望めないことを説明した。ところが、時系列データを扱い、再帰的構造を持つネットワークを重ねる場合には性能の向上が望める場合があることが発見されている[16, 17]。

文献[16, 17]用いたデータセットを図4に示す。図4のデータセットは、(64ピクセル)×(64ピクセル)×(3チャンネル)の連続値によってなる画像26枚から構成されている。RBMによって抽出された特徴量を用いても時系列の学習が困難であるが、IL-RBMによる特徴抽出を用いることで容易に学習できることを示している。

素子が動的に増加するモデルから再帰構造をもつモデルへの情報の流れは、図1で示した歯状回からCA3への神経投射とよく対応する。

## 4. おわりに

海馬の構造的特徴と共通点をもつRBMの拡張モデルがもつ機能について紹介した。歯状回の新生ニューロンと対応付けて議論したモデルは、既学習情報を失わない逐次学習や、高速な学習、抽象度の異なる表現の獲得が可能であることを示唆した。一方強い再帰的な神経投射があるCA3に対応付けて議論したモデルは、時系列処理や連想記憶、作業記憶の機能を持つことを示唆した。

さらに、海馬周辺の神経接続を参考にモデルを組み合わせることでより優れた性能が得られる例を示した。

## 参考文献

- [1] 佐藤 直行:海馬から大脳へ: 記憶の計算モデル, 人工知能学会全国大会, (2015)
- [2] Altman, Joseph: Autoradiographic investigation of cell proliferation in the brains of rats and cats., *The Anatomical Record*, Vol. 145, No.4 pp. 573-591, (1963)
- [3] Schmidt-Hieber, Christoph, Peter Jonas, and Josef Bischofberger: Enhanced synaptic plasticity in newly generated granule cells of the adult hippocampus., *Nature*, Vol. 429, No. 6988, pp. 184-187, (2004)
- [4] Clelland, C. D., et al.: A functional role for adult hippocampal neurogenesis in spatial pattern separation., *Science*, Vol. 325 No.5937 pp. 210-213, (2009)
- [5] Kitamura, Takashi, et al.: Adult neurogenesis modulates the hippocampus-dependent period of associative fear memory., *Cell*, Vol. 139 No.4, pp. 814-827, (2009)
- [6] MacVicar, Brian A., and F. Edward Dudek: Local synaptic circuits in rat hippocampus: interactions between pyramidal cells., *Brain research*, Vol.184 No.1 pp. 220-223, (1980)
- [7] Marc-Alexandre Cote, and Hugo Larochelle: An Infinite Restricted Boltzmann Machine, *Neural Computation*, Vol.28, No. 7, pp. 1265-1288, (2016)
- [8] 大澤正彦, 萩原将文.: RBMにおける未学習データ検出法の提案と追加学習への応用, *信学技報*, Vol. 114, No. 259, pp.283-288, (2015).
- [9] Platt, J.: A resource allocating network for function interpolation, *Neural computation*, Vol. 3, No. 2, pp. 213-225, (1991)
- [10] Zhou, G., Sohn, K., and Lee, H.: Online incremental feature learning with denoising autoencoders, in *International Conference on Artificial Intelligence and Statistics*, pp. 1453-1461, (2012)
- [11] Carpenter, G. A., and Grossberg, S.: The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network, *Computer*, Vol. 21, pp. 77-88, (1999)
- [12] Sutskever, Ilya, and Geoffrey E. Hinton.: Learning multilevel distributed representations for high-dimensional sequences., *International Conference on Artificial Intelligence and Statistics*, (2007)
- [13] Sutskever, Ilya, Geoffrey E. Hinton, and Graham W. Taylor.: The recurrent temporal restricted boltzmann machine. *Advances in Neural Information Processing Systems*, (2009)
- [14] N Boulanger-Lewandowski, Y. Bengio and P. Vincent.: Modeling Temporal Dependencies in High-Dimensional Sequences: Application to Polyphonic Music Generation and Transcription, in *Proceedings of the 29th International Conference on Machine Learning*, (2012).
- [15] 大澤正彦, 萩原将文: 海馬-大脳新皮質系に着目した深層学習による記憶モデルの提案, 人工知能学会全国大会, (2015)
- [16] Yu Yamagishi, Masahiko Osawa, Masafumi Hagiwara.: A Learning Method for Echo State Networks Using RBM, *International Symposium on Advanced Intelligent Systems*, (2015)
- [17] Masahiko Osawa, Michita Imai, and Hiroshi Yamakawa.: An Implementation of Working Memory Using Stacked Half Restricted Boltzmann Machine: Toward to Restricted Boltzmann Machine-Based Cognitive Architecture., *The 23rd International Conference on Neural Information Processing*, (2016)
- [18] Marr, D.: Simple memory: a theory for archicortex., *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, Vol. 262, No. 841, pp. 23-81, (1971)