

前頭前野の Accumulator モデルを利用した、 階層型脱抑制システムの提案

Proposal of Hierarchical Derepression Model using Prefrontal Accumulator Model

芦原佑太 *1*4
Ashihara Yuta

大澤正彦 *2
大澤正彦

島田大樹 *3
Daiki Shimada

栗原 聡 *1
Kurihara Satoshi

今井 倫太 *2
Imai Mitita

*1電気通信大学 大学院情報理工学研究科 情報学専攻

The University of Electro-Communications, Graduate School of Informatics and Engineering, Department of Informatics

*2慶應義塾大学 大学院 理工学研究科

Graduate School of Science and Technology, Keio University

*3法政大学 大学院理工学研究科

Graduate School of Science and Engineering, Hosei University

*4株式会社クロスコンパス・インテリジェンス

XCompass Intelligence Ltd.

Abstract: According to past findings, prefrontal area is thought to have the function of derepression and arbitrating many brain regions, and reported that it is involved in the initiation of voluntary exercise. Some specific cells acting as suppressor, and this kind of cells exist in prefrontal area. In our work, with reference to the findings reported about the prefrontal area, propose a hierarchical arbitration method using accumulator model.

1. はじめに

前頭前野において、神経細胞の抑制・脱抑制を行うような、Accumulator として活動する神経細胞が、自発的な運動の開始に関与しているという主張がされている [1]。本稿ではこれまで前頭前野に関して報告されている知見を参考に、Accumulator モデルを用いたアーキテクチャを構築し、階層型に配置された機械学習器の抑制・脱抑制を行う方法について提案する。Damasio ら [1] によれば、前頭前野を損傷した患者の観察の結果、適応的な社会活動に支障をきたす行動障害が発現し、日常生活で起こる対人場面や、問題解決を行う場面などで、トラブルを引き起こしやすくなるという現象が報告されている。つまり、前頭前野の損傷は、利他的な行動など、目標に対して適切な行動を行う際の順序付けなどがうまくいかない現象を引き起こすと考えられている。

一方で、近年扱われることの多い、DeepLearning を用いた行動制御においては、DeepLearning が出力する値を行動として変換し、即時的に実行する制御型が多く用いられていることから、モデルは特定の時間間隔で何かしらの行動を出力することになり、無駄な行動も多く見られる。そこで、本研究では、前頭前野を参考に、行動選択時に抑制・脱抑制の機能を導入することで、無駄な行動を少なくするような、モデルの提案を目指す。具体的には、Accumulator を用いた複数の機械学習器を調停する階層型モデルを提案し、分類問題、強化学習に適用することのできるモデルを構築する。本稿では、次節で関連研究について紹介し、3 節ではモデルの概要について述べる。4 節では、モデルを使った実験について述べ、最後に 5 節でまとめを述べる。

2. 関連研究

Schurger ら [3] は、動物の自発的な運動の開始は Accumulator モデルを用いてモデル化が可能であるとしており、Soon-

連絡先: 芦原佑太, 電気通信大学大学院情報理工学研究科,
〒 182-8585 東京都調布市調布ケ丘 1 丁目 5-1, E-mail:
y.ashi@ni.is.uec.ac.jp

Haynes ら [4] は、自発的な運動選択の開始部位として、前頭前野が関与していると示唆している。また、意思決定を行う認知アーキテクチャの既存研究として、Brooks[5] によるサブサンクションアーキテクチャがある。サブサンクションアーキテクチャは、実世界的なロボットの行動を可能とするために考案された階層型アーキテクチャと言われている。サブサンクションアーキテクチャは下位から上位に向けて優先順位が高くなり、上位の行動が選択される時には、下位の行動は抑制される。サブサンクションアーキテクチャの階層性は、時には即時的な行動、時には慎重な行動の選択を実現する手法の一つである。また、特長として、階層のどこかが故障してしまっている場合にも、他の階層が行動を取ることで、システムが破綻しにくいという点もある。これらの知見から、本研究では、前頭前野の Accumulator モデルを採用した階層型アーキテクチャを提案する。

3. 提案モデル

本節では、提案モデルの概要について述べる。提案するモデルは、複数の機械学習器を階層的に調停するモデル (提案モデル 1) 及び、強化学習系に応用するための階層型モデル (提案モデル 2) である。提案モデル 1 では、各学習器を Convolutional Neural Network で構成しており、各学習器は下位の学習器から上位の学習器に向かって、精度の高い順、あるいは低い順に配置される。

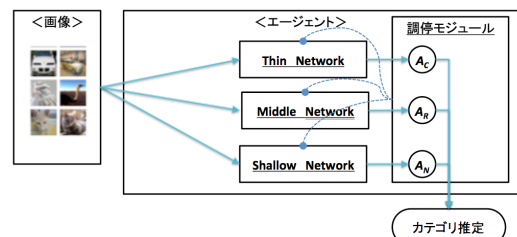


図 1: 提案モデル 1

提案モデル1では、各学習器がそれぞれ画像を受け取り、その画像のカテゴリを推定する。下位のモデルの出力値は、そのまま証拠値として蓄積され、Accumulatorの閾値の達していない場合は、一つ上の階層のモデルからの出力をさらに証拠値として蓄積していく。蓄積された証拠が一定の閾値を越えると、閾値を超えたカテゴリが結果として出力される。今回は閾値 α は $[0, 2.0]$ で設定している。

強化学習系に応用するモデルについては、複数のモジュールを調停する仕組みになっている。各モジュールは提案モデル1のように全ての階層に機会学習器を配置せず、下位のモジュールはルールベース(静止モジュール、反射モジュール)となっており、上位のモジュールがDeep Q Network(認識モジュール)を配置するモデルとなっている。

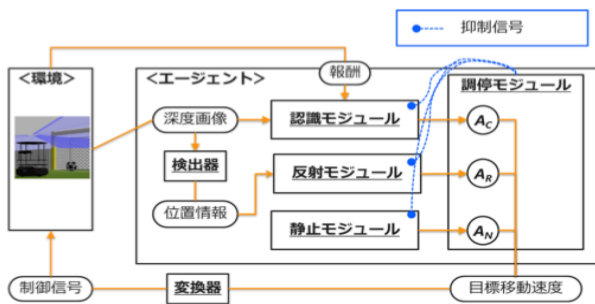


図 2: 提案モデル 2

強化学習系では、環境からの入力を反射モジュールと認識モジュールが受け、それぞれのモジュールが行動を選択する。選択された行動が Accumulator に送られ、証拠として蓄積し、蓄積された証拠が閾値を越えると、その行動が出力される仕組みとなっている。また、静止モジュールを下位に配置することにより、反射モジュールと認識モジュールが閾値を越えるような動作出力をしていない間は、静止し、周辺の情報探索のみを行うことができる。

4. 提案モデルを用いた実験

本節では、提案モデルを用いて行った実験についての結果を示す。

4.1 階層型モデルの調停実験

4.1.1 データセット

本研究の実験で使用するデータセットは、[8]で公開されている CIFAR-10 を用いる。CIFAR-10 は、10 カテゴリの画像が用意されており(図 3. 参照)、各カテゴリ毎に画像が 6000 枚用意されている。実験では、学習用に 50000 枚、検証用に 10000 枚を使用した。

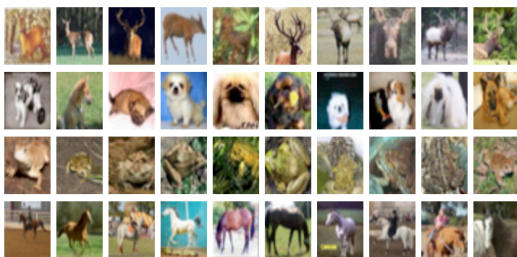


図 3: CIFAR-10 の画像の一部

4.1.2 学習器の構成

前節の図 1. 図 2. にあるように、階層型に配置する 3 つの学習器はそれぞれ、Shallow Network, Middle Network, Thin Network としており、各学習器は表 1. のレイヤー構成で実験を行った。

表 1: 実験でのレイヤー構成

Network 名	Shallow	Middle	Thin
畳込み層の数	4	4	5
活性化関数	relu	relu	elu

4.1.3 実験結果

本研究では、3 つのモデルパターンを用意し、実験を行った。まず、提案モデルとの比較対象として、Shallow Network, Middle Network, Thin Network の 3 種類の Network を単体で学習した。Network を単体で行った際の分類精度については、それぞれ表 3. の”Shallow only”, ”Middle only”, ”Thin only”に記載している。次に、Zhang ら [6] が行った、Ensemble 法について 3 種類の Network をベース (Ensemble1) に、Softmax 関数がある場合、ない場合、Network を 2 種類に厳選して実験した場合 (Ensemble2), 減衰パラメータ (decay) を導入し、Thin Middle Shallow の順に減衰率を大きくする条件 (Ensemble3) と、逆の Shallow Middle Thin の順に減衰率を大きくする条件 (Ensemble4) において学習した。提案モデルについては、下位の学習器から上位の学習器の順番が、Shallow Thin で行った場合 (提案モデル 1), また、学習器の順番を、Thin Shallow の順で行った場合 (提案モデル 2) について実験を行い、Ensemble 学習と同様に、decay パラメータを導入し、Thin Middle Shallow の順に減衰率を大きくする条件 (提案モデル 3) と、逆の Shallow Middle Thin の順に減衰率を大きくする条件 (提案モデル 4) において学習した。各モデルの実験条件は表 2. についてまとめた。

表 2: 実験条件のまとめ

Network 名	実験条件
Shallow only	Shallow Network 単独での学習
Middle only	Middle Network 単独での学習
Thin only	Thin Network 単独での学習
Ensemble1	Ave Ensemble を用いた学習
Ensemble2	Shallow, Middle で Ensemble 学習
Ensemble3	decay を導入した Ensemble 学習
Ensemble4	decay を導入した Ensemble 学習
提案モデル 1	Shallow Thin の順の階層で学習
提案モデル 2	Thin Shallow の順の階層で学習
提案モデル 3	decay を導入した学習
提案モデル 4	decay を導入した学習

実験結果については、表 3. にまとめた。実験の結果、各学習器を単独で学習する場合に比べて、提案モデルは高い精度を記録していることを確認することができた。また、Zhang ら [?] の Ensemble 法との比較についても、decay を導入した提案モデル (提案モデル 3, 提案モデル 4) は同等に近い精度が出ている。今後、Accumulator の閾値の設定をどのように設定すれば良いか、証拠の累積値とその減衰値について検討を行

うことで、より精度を上げていくことが可能であると考えられる。

表 3: 実験でのレイヤー構成

Network 名	Accuracy
Shallow only	0.7439
Middle only	0.7613
Thin only	0.7131
Ensemble1	0.7947
Ensemble2	0.7833
Ensemble3	0.7880
Ensemble4	0.7881
提案モデル 1	0.7882
提案モデル 2	0.7886
提案モデル 3	0.7911
提案モデル 4	0.7881

4.2 強化学習系への応用実験

4.2.1 環境設定

本実験での環境は、Gazebo+ROS を使った、サッカー用のフィールドで、タートルボットに PK を行うタスク環境を設計した。タートルボットは開始時に特定のエリア内にランダムに置かれ、ボールをゴールに入れるタスクを行うものとした。



図 4: タスク環境

ボールがゴールに近づけば近づくほど報酬が与えられるものとして、ゴールに至るまでのボールのゴールまでの垂直距離を 0 1.0 点の範囲で報酬とし、ゴールできた時点で 1.0 点を与え、10 秒以内にゴールできなかった場合は-1.0 点を与えるような環境を設計した。タートルボットの行動サイクル毎に行うことのできる行動は、前進、後退、回転(右回りと左回り)、静止の合計 5 つのアクションの中から選ぶものとした。

4.2.2 実験結果

DQN 単体で学習を行った場合と、提案モデル 2 を使って学習の経過を見た場合で、タスクがどれくらい出来るようになったかを確認したところ、DQN は 13 時間学習を行っても、特定の位置からのスタート以外ではゴールできない状態となっており、まだまだ汎化するには多くの思考が必要とされていたが、提案モデルでは、同じ 13 時間の学習でも、ゴールする回数が増えており、DQN 単体での学習に比べて、提案モデル 2 では学習が早く進んでいることを観測することができた。本実験の定量化については、今後の課題としたい。

5. 終わりに

本稿では、前頭前野の Accumulator ニューロンを用いた複数の学習器を調停するモデルについて提案した。提案モデル 1 では、学習器を階層に配置し、特定の閾値に達するまで、下位の学習器から順に出力を証拠値として蓄積する仕組みでカテゴリ推定を行い、学習器が単独で学習を行った結果よりも高い精度でカテゴリ推定が行える結果となった。しかし、Zhang ら [?] の Ensemble 法で実装した結果と比べて、優位であるとは言えず、今後は、閾値 α の調整、証拠の減衰率の調整、階層モデルの配置順を検討するなど、様々な工夫をする必要があることも明らかになった。また、強化学習系に提案モデル 2 を適用してみたところ、DQN を単体で扱うよりも、短い学習時間でタスクの学習ができることを期待できる。今後の課題として、別のタスクへ提案モデルを応用していくことで、階層型の抑制・脱抑制を行う有用性を示していきたい。

参考文献

- [1] Damasio, A. R.: Descartes' error Emotion, reason, and the human brain. *New York Putnam.*(1994)
- [2] F. Chersi and N. Burgess.: The cognitive architecture of spatial navigation Hippocampal and striatal contributions, *Neuron*, vol.88, no.1, pp.64-77 (2015)
- [3] Schurger A, Sitt JD, Dehaene S.: An accumulator model for spontaneous neural activity prior to self-initiated movement, *Proc Natl Acad Sci USA*, pp.2904-2913 (2012)
- [4] John-Dylan H, Chun S, Marcel B, Hans-Jochen H.: Unconscious determinants of free decisions in the human brain, *Nature Neuroscience*, vol.11, pp.543-545 (2008)
- [5] Rodney A Brooks.: A Robust Layered Control System For A Mobile Robot, *IEEE Journal of Robotics and Automation*, vol.RA-2, pp.14-23 (1986)
- [6] Xiaohui Zhang, Daniel Povey, Sanjeev Khudanpur.: A diversity-penalizing ensemble training method for deep learning, in *Proceedings of the 16th Annual Conference of International Speech Communication Association*, (2015)
- [7] Bolei Zhou, Agata Lapedriza, Jianxiong Xiao, Antonio Torralba, Aude Oliva.: Learning Deep Features for Scene Recognition using Places Database, *Neural Information Processing Systems* (2014)
- [8] CIFAR-10 and CIFAR-100 datasets
<https://www.cs.toronto.edu/~kriz/cifar.html> (アクセス日: 2017/2/21)