

## 「人工知能と脳科学」領域の5年間とこれから

新学術領域研究「人工知能と脳科学」領域代表 / 沖縄科学技術大学院大学 神経計算ユニット 教授 銅谷 賢治

私たちの新学術領域研究「人工知能と脳科学の対照と融合」は、「それぞれの研究の高度化のなかで乖離して行った人工知能研究と脳科学研究を再び結びつけ、両者の最新の知見の学び合いから新たな研究ターゲットを探り、そこから新たな学習アルゴリズムの開発や脳機構の解明を導く」ということを目標に、2016年度にスタートし2021年3月に終了を迎えました。計画研究11課題、公募研究前期18課題、後期20課題の連携のもと、人工知能技術や計算理論により脳機能を解明する「AIから脳」、脳科学の知見を次世代の人工知能の設計開発に活かす「脳からAI」、さらに新たな研究コミュニティを形成し人材育成を行う「AI脳融合」という3つの方向で研究活動を展開して来ました。この約5年の取り組みの中で、新たな脳科学的発見や人工知能技術の開発が進んだだけでなく、両分野の研究者が「知能はどう生まれるのか」という共通の関心に向け頭を寄せ合い取り組むネットワークが動きだしたこと、またAIと脳科学の双方の知見や技術をもとに将来の科学技術を担う若い研究者たちの姿を見るにつけ、この領域を立ち上げて良かったと改めて感じています。

「AIから脳」という方向では、強化学習やベイズ推定の理論をもとに計画した実験やデータ解析によって、ドーパミン細胞が行動の選択肢の評価から行動選択に関して大脳皮質よりも早く応答することや、セロトニンが報酬予測の事前確率や、モデルフリーとモデルベースの意思決定のバランスを制御するといった新たな発見が得られました。

「脳からAI」という方向では、ヒトや動物の脳の階層並列学習制御機構をもとに、人型ロボットの運動スキルの制御方式や、報酬と罰からの並列的な強化学習アルゴリズムを開発し、シミュレーションやロボット実験でその有効

性を実証することができました。さらに脳全体を確率生成モデルのネットワークとして捉えて、それを汎用的な人工知能の設計原理とするという構想を打ち出し、それを実装するためのツールNeuro-SERKETを開発、公開しました。



「AI脳融合」の面では、両分野の研究者を集めたワークショップやシンポジウムを多数開催するとともに、理研CBSと共催のサマーコース、東大IRCNと共催のチュートリアルコースを開催するなど、若手の育成に力を入れて来ました。2020年10月に開催したInternational Symposium on Artificial Intelligence and Brain Scienceは両分野の先端を走る研究者を講師に迎え、オンラインながら参加登録者が1,800名以上にのぼる反響を呼び、その成果をもとにNeural Networks誌の特集号が出版されています。

この領域から生まれた国際的ネットワークと若手研究者は、AIと脳科学を融合する次世代の研究を大いに発展させてくれることでしょう。それを可能にしてくれた、科研費新学術領域研究制度とその審査員やアドバイザー、学術調査官や事務局の皆様、またこのニュースレターを読まれ私たちの研究を叱咤激励して下さった皆様に心より感謝いたします。



人工知能と脳科学サマースクール 2017

### 多階層表現学習の数理基盤と神経機構の解明

研究代表者：銅谷 賢治（沖縄科学技術大学院大学 神経計算ユニット 教授）

近年の人工知能のブレークスルーは、ディープラーニングによる多層神経回路の表現学習と、強化学習による状態や行動の価値評価を結びつけることによって実現しました。私たちは、ディープラーニングを強化学習により確実に効率よく適用するための手法の開発と、脳の柔軟な強化学習を実現している神経回路や物質の機構を解明するための生物実験を両輪として研究を進めて来ました。

1) 多階層表現学習の数理基盤：ディープラーニングと強化学習は、これから得られる報酬の予測や、環境の現在の状態の推定や未来の状態の予測など、様々な形で組み合わせることができます。私たちは、複雑な環境の予測モデルを強化学習に用いる場合の理論的な課題とその解決策 (Parmas et al., ICML2018)、学習の安定性を保証しながらデータ効率を改善するアルゴリズムと数理解析 (Kozuno et al., AISTATS2019)、ループを持つニューラルネットによる隠れ状態の推定を用いた強化学習 (Han et al., ICLR2020) など、新たな手法を開発し、AI分野の先端的な国際会議で発表を行いました。

2) 多階層表現学習の神経機構：脳の強化学習では、神経伝達物質のドーパミンと、その入力を強く受ける大脳基底核が重要な役割を果たすと考えられています。大脳基底核は複数の神経核とそれらの中の様々な種類の神経細胞から構成されています。私たちはそれらの中でもドーパミン細胞に直接出力を送る種類の神経細胞の活動を、最新の内視鏡型の顕微鏡と、細胞種選択的なカルシウム蛍光イメージングにより捉えることに成功しました。解析の結果これらの細胞は、マウスが特定の匂いから報酬や罰を予測する学習が進むにつれて、予測される報酬や罰に応じて活動をするようになることを明らかにしました (Yoshizawa et al., eNeuro, 2018)。

3) 脳の学習の制御機構：脳の学習や行動制御は、大脳皮質や大脳基底核などの様々な機能モジュールが柔軟に連携し合う

ことで実現されています。私たちは、強化学習においてどれだけ先の報酬を考慮すべきかという問題に関して、神経伝達物質のセロトニンが、その脳の幅広い領野への出力によってバランス制御を行っているという仮説をもとに、その検証実験を進めて来ました。オプトジェネティクスという手法で、セロトニン神経細胞を選択的に光で刺激すると、マウスはエサの報酬をより長く待てるようになります (Miyazaki et al., Current Biology, 2014)。私たちはさらに、その効果は報酬が確実に得られるけれどもそのタイミングが不確かな場合に特に強いことを明らかにしました (Miyazaki et al., Nature Communications, 2018)。この現象は、マウスは報酬の待ち時間の予測モデルをもとに待ち続けるか諦めるかの意思決定を行い、セロトニンは報酬が得られる事前確率を制御するという数理モデルによって再現することができました。また、セロトニン神経の異なる投射先を光刺激することで、前頭葉の異なる部位が報酬予測に関わることを明らかにしました (Miyazaki et al. Science Advances, 2020)。これらの研究により、脳の柔軟な学習の仕組みの一端を明らかにすることができましたが、今後さらにこれらの知見を脳のように柔軟に学習するAIシステムの開発につなげることを目指していきます。

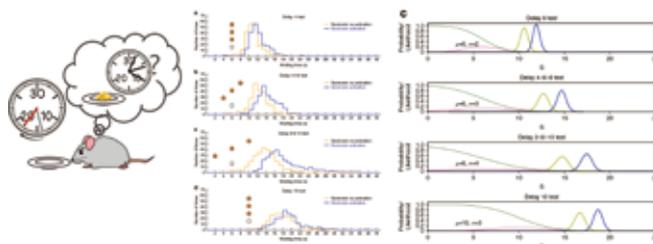


図 セロトニン細胞の光刺激による報酬待機行動の促進は、報酬のタイミングが不確かな時ほど強くなる。マウスは報酬のタイミングの予測モデルをもとにベイズ推定により報酬の有無を判断し、セロトニン刺激は報酬が得られる事前確率を高めるとすると、実験結果を再現することができる (Miyazaki et al., Nature Communications, 2018)。

#### 主な研究業績：

- 1) Parmas P, Rasmussen CE, Peters J, Doya K (2018). PIPPS: Flexible model-based policy search robust to the curse of chaos. 35th International Conference on Machine Learning (ICML 2018). <http://proceedings.mlr.press/v80/parmas18a/parmas18a.pdf>
- 2) Yoshizawa T, Ito M, Doya K (2018). Reward-predictive neural activities in striatal striosome compartments. eNeuro, 5, e0367-17.2018. <https://doi.org/10.1523/ENEURO.0367-17.2018>
- 3) Miyazaki K, Miyazaki KW, Sivori G, Yamanaka A, Tanaka KF, Doya K (2020). Serotonergic projections to the orbitofrontal and medial prefrontal cortices differentially modulate waiting for future rewards. Science Advances, 6, eabc7246. <https://doi.org/10.1126/sciadv.abc7246>

## コンフリクトコストに対する調和・不調和情報シーケンス効果の神経基盤の研究

研究代表者：田中 啓治（理化学研究所脳神経科学研究センター 特別顧問）

私たちは無意識のうちに認知制御のレベルを課題遂行に必要最低限近くに留めます。認知制御のリソースが限られているからです。具体的には、うまく行っている限りは制御（課題遂行への集中）のレベルを徐々に緩める、しかし失敗の危険を察知すると制御のレベルを上げます。すなわち認知制御のレベルは動的に調節されています。この認知制御の動的調節は前頭前野の認知制御メカニズムを解きほぐすための良い具体的問題です。これまで認知制御の動的調節は主に競合課題を用いて調べられてきました。例えば色を表す単語の文字インク色を答えるスループ課題では、自然に起こる文字を読む反応が課題で指示されたインク色を答える反応と競合します。文字の色と単語の意味が一致しない不調和試行では調和試行に比べて反応時間が長く（競合コスト）、また不調和試行に続く試行での競合コストは、調和試行に続く試行での競合コストに比べて、小さい（調和シーケンス効果）ことが知られています。調和シーケンス効果はこれまで「不調和試行での反応競合の経験の結果、次試行では認知制御（インク色への集中）のレベルが上がる」という競合モニター仮説により説明されてきました。しかし最近になって競合モニター仮説では説明できないいろいろな行動上の結果、また機能ブロック法による結果が報告されています。本研究では、前試行で励起された反応セットの影響が次試行まで残るといった受動的な過程とそれへの対応の過程を含めたより総合的な考えを提案することを目指しました。経頭蓋直流電気刺激（tDCS, transcranial direct current stimulation）による局所脳活性化操作の課題遂行への影響を調べる研究とfMRIでの課題関連脳活動の測定を行なう研究を行いました。ここでは紙面の都合でfMRI研究の結果だけを紹介します。

色のついた線分と矢印の重ね合わせを刺激とするサイモン課題を用いました。まず、ある恣意的対応規則によって色が指定する方向のボタンを押す課題を被験者に訓練しました。次に、色が指定する方向と同じ（一致条件）または反対（不一致条件）の方向を指す矢印、または矢頭なしの線分（中立条件）を色刺激の上に置き、矢印が自然に引き起こす反応が色一方向課題が指定する反応に干渉するようにしました（図1）。そしてfMRIにより測定した脳活動を脳フラットマップにマップし、フラットマップ上で円形のROIを動かし、ROI内のボクセル集団の活動パターンから課題情報のデコードを試みしました（MVPA, multi-voxel pattern analysis）。その結果、右前頭

眼窩皮質外側部の活動パターンが前試行の条件と現試行の条件を表現すること（図2）、さらに被験者間の比較において、競合コストが前試行の条件に依存する依存度の大きさが、眼窩皮質における前試行条件表現の強さに負相関し、現試行条件表現の強さに正相関することが明らかになりました。

これらの結果は、前頭眼窩皮質外側部が前試行で活性化したタスクセットの数を表現し、これを現試行の認知制御に用いていることを示しました。従来、前頭眼窩野の働きは刺激—報酬連合学習であると考えられてきました。今回の結果を踏まえて、私たちは「前頭眼窩皮質外側部は、内的な課題遂行状態（task state）を表現することにより、認知制御に貢献する」という新しい考えを提案します。

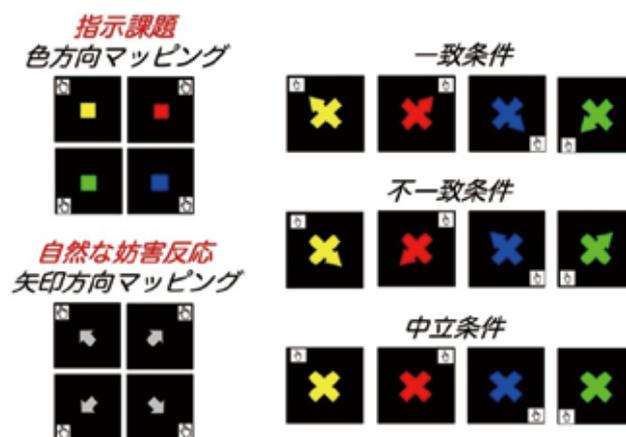


図1 fMRI実験に用いたサイモン課題

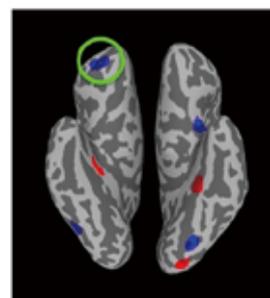


図2 右前頭眼窩皮質外側部の活動パターンによる前試行の条件の表現

### 主な研究業績：

- 1) Li N, Wang Y, Jing F, Zha R, Wei Z, Yang LZ, Gneg X, Tanaka K, Zhang X (2020). A role of the lateral prefrontal cortex in the congruency sequence effect revealed by transcranial direct current stimulation. *Psychophysiology* 2021;00:e13784. <https://onlinelibrary.wiley.com/doi/10.1111/psyp.13784>

## ディープラーニングと記号処理の融合による予測性の向上に関する研究

研究代表者：松尾 豊（東京大学 工学系研究科 教授）

本研究は、深層学習の進展のさらに先を見据え、深層学習と記号処理の融合に挑戦しようというものです。こうした問題意識は、最近では深層学習の学術コミュニティでもよく理解されており、なかでも、2019年にニューラルネットワークの国際会議 NeurIPSにおいて、Yoshua Bengio氏が行った「システム1の深層学習からシステム2の深層学習へ」というタイトルの招待講演が有名です。システム1、2というのは心理学者ダニエル・カーネマンの用語ですが、システム1は速くて直感的な思考（現在の深層学習に該当）、システム2は遅くて論理的・言語的な思考（記号処理に該当）を指します。本研究は、その問題意識を2016年という早い時期から捉えていたものです。

本研究では、深層学習と記号処理を融合するために、2つのテーマを設定しました。1点目は、深層強化学習を発展させることです。当時、さまざまなコンピュータゲームを学習によって上達することのできる Deep Q Network (DQN) がよく用いられていました。これに記号処理を組み込むことによって進展させようということが目的です。

2点目は、文章からの画像の生成モデルを用いて、推論に該当するような処理を実現しようというものです。研究提案当時

は、Generative Adversarial Network (GAN) などの深層生成モデルが注目され、さまざまな改良の手法が出始めた時期でした。これを文章の意味理解に活用しようというものでした。

研究の過程で紆余曲折がありましたが、結果的には、トップ国際会議の論文4本を含め、多くの研究成果を出すことができました。例えば、情報処理学会の論文賞の受賞にもつながった研究では、マルチモーダルの深層生成モデルに関する手法を提案しました。従来、マルチモーダルなデータにおいて、一部のモーダルが欠損した場合に、うまく生成を行うことが難しかった問題を解決するもので、手法としての汎用性が高いものです。また、ICLR2021においては発表した論文は、実際のロボット等を使った強化学習においては、デプロイする（つまりモデルを実機に送り実機を動かしてデータを取る）ことのコストが高いという問題がありますが、世界モデルを用いるなどのモデルベースな手法は、デプロイの数を減らすことができます。そうして、そこで、デプロイの数を減らしながら、同時に学習の精度を上げていく方法を提案し、新規性の高いものとなりました。

本研究で得られた成果をベースに、今後さらに、記号処理と深層学習の融合について研究を進めていきたいと思っています。

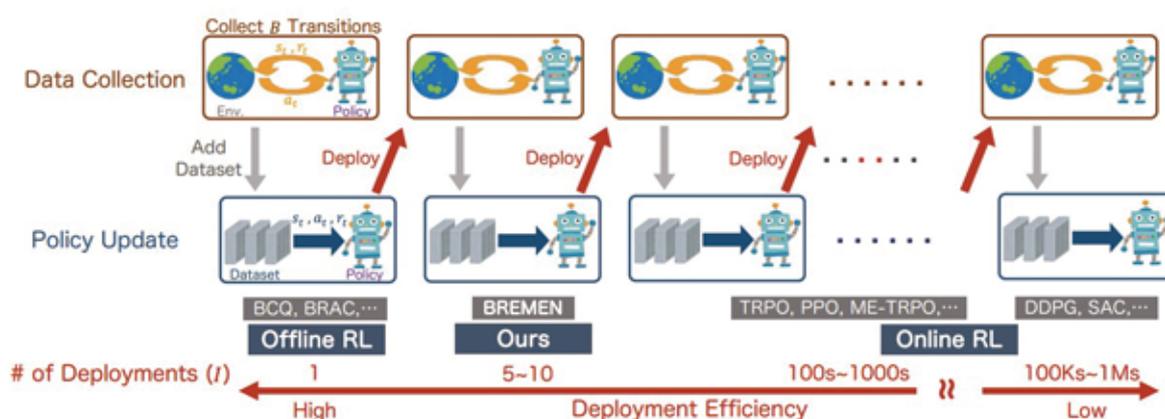


図 デプロイ効率の高い強化学習手法

### 主な研究業績：

- 1) 銅谷賢治, 松尾豊 (2019). 人工知能と脳科学の現在とこれから. Brain and Nerve, 71, 649-655. <https://doi.org/10.11477/mf.1416201337>
- 2) 鈴木雅大, 松尾豊 (2018). 異なるモダリティ間の双方向生成のための深層生成モデル. 情報処理学会論文誌, Vol. 59, No. 3 (2018) (2018年度 情報処理学会論文賞)
- 3) Matsushima T, Furuta H, Matsuo Y, Nachum O, and Gu S (2021). Deployment-Efficient Reinforcement Learning via Model-Based Offline Optimization. International Conference on Learning Representations 2021 (ICLR2021).

## 前頭前野活動の網羅的計測と情報表現解読法の開発

研究代表者：宇賀 貴紀（山梨大学大学院 総合研究部 医学域統合生理学 教授）

環境の変化に応じて多様な意思決定をする能力は、霊長類で特に発達した本質的な脳機能です。この機能に前頭前野が深く関わっていることが、これまでに明らかにされています。しかし、前頭前野の情報表現は複雑であり、前頭前野で行われている計算が、どのような大脳皮質ネットワークで実現されているのかを理解するのは非常に困難です。そこで本研究では、2つのルールに基づいて、判断に用いる情報を切り替えるタスクスイッチ課題中に、前頭前野の複数領域から皮質脳波（ECoG）を計測し、神経活動を網羅的に取得し、前頭前野で表現されている情報を解読することを試みました。

新たに設計、作成したECoG電極を前頭葉眼窩面／大脳半球間裂面／背外側面に留置し、タスクスイッチ課題中のECoG信号を計測しました。タスクスイッチ課題では、最初にルールに

関する手がかりが呈示されますが、その情報をどの領野で検出できるのか、サポートベクターマシンを用いて解読しました。その結果、眼窩前頭前野と背外側前頭前野のECoG信号から、ルールの情報を解読できました。さらに、眼窩前頭前野のECoG信号の方が、背外側前頭前野のECoG信号よりも早い時間にルールを反映することがわかりました。このことから、どの情報を判断に利用すべきかというルールに関する情報は、一旦眼窩前頭前野で処理されたあとに、背外側前頭前野に伝達されると考えられます。

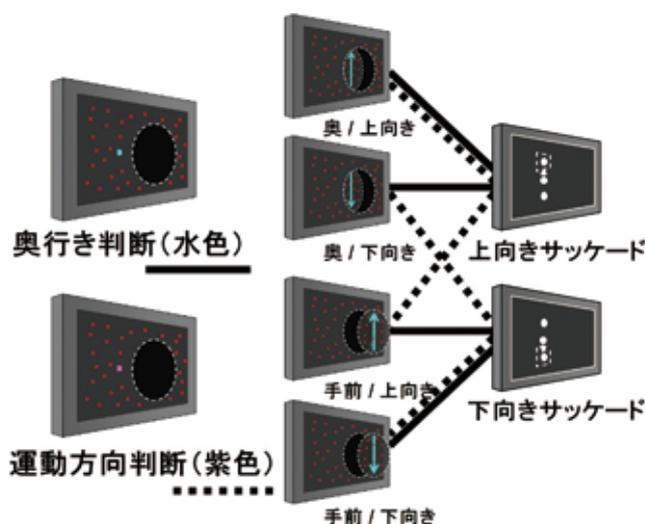


図 タスクスイッチ課題では、運動方向および奥行き判断課題を課す。試行ごとに注視点の色を変え、ランダムに2つの判断課題のどちらかを行なうように指示する（左パネル）。運動方向判断の場合、ドットが上向きに動いていたら目を上に向け、ドットが下向きに動いていたら目を下に向ける。奥行き判断の場合、注視点よりもドットが奥にあれば目を上に向け、ドットが手前にあれば目を下に向ける（中パネルと右パネル）。

### 主な研究業績：

- 1) Sasaki R, Kumano H, Mitani A, Suda Y, Uka T (in press). Task-specific employment of sensory signals underlies rapid task switching. Cerebral cortex.

## 神経信号からネットワーク構造を推定し、そこに発現する活動パターンを予測する

研究代表者：篠本 滋

近年の神経活動計測技術の急速な発展により、多数の神経信号を同時計測することが可能になってきています。私たちは複数ニューロンの神経スパイク列の解析手法の開発に取り組みました。本研究課題ではマルチ同時計測データから神経ネットワークを推定することを目標にあげました。

神経スパイク信号は神経結合を通して他の神経に伝わりますので、神経発火信号の相互相関をみればニューロン間の結合を推定することができる、ということは50年以上前から提唱され

ていました。しかし各ニューロン対はそれ以外の多くの神経細胞の影響を受けているため、注目するニューロン間にいくつかの同期スパイク発生が計測されたからといってそれでシナプス結合があると結論づけることはできません。この難問を解決するにはこの新学術の公募班2年間を含めて6年の歳月を要しました。私たちは後半の公募には採択されませんでした。最終的には一般化線形モデルを用いた結合推定法を完成させ、さらに人工知能を用いた推定法も提案しました。

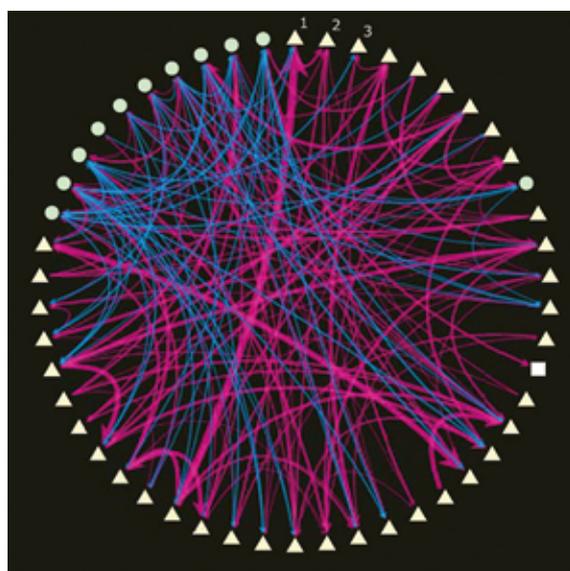


図1 神経信号から推定した単シナプス結合。興奮性はマゼンタ、抑制性はシアンで表し、結合の強さは線の太さで表した。

### 主な研究業績：(期間外)

- 1) Kobayashi R, Kurita S, Kurth A, Kitano K, Mizuseki K, Diesmann M, Richmond BJ, and Shinomoto S (2019) Reconstructing neuronal circuitry from parallel spike trains. Nature Communications 10:4468  
<https://www.nature.com/articles/s41467-019-12225-2>
- 2) Endo D, Kobayashi R, Bartolo R, Averbek BB, Sugase-Miyamoto Y, Hayashi K, Kawano K, Richmond BJ, and Shinomoto S (2021) A convolutional neural network for estimating synaptic connectivity from spike trains. Scientific Reports 11:12087  
<https://www.nature.com/articles/s41598-021-91244-w>

## 人工知能と神経基盤の相互参照アプローチによる視覚-価値変換機構の解明

研究代表者：近添 淳一（株式会社アラヤ 脳事業研究開発室 チームリーダー）

この研究は、人工神経回路と脳内の情報表現を比較することにより、未知の脳内情報処理過程を明らかにしました。まず、油絵の美的価値を推定する人工神経回路を作成し、油絵の評価を行う際の脳活動を機能的MRIによって計測した上で、これらの対応を、機械学習の手法を使って同定しました。その結果、人工神経回路の入力（油絵画像）に近い低次の階層は初期視覚野と対応していたのに対し、出力（美的価値）に近い高次の階層は、いわゆる default mode network と対応していることが明らかにされました（図1）。

この研究は、しばしばブラックボックスであるといわれる人工神経回路の情報処理過程を脳と対応付けることで視覚化できるこ

とを示した上で、神経科学研究の領域のみでなく、人工知能研究の領域においても有意義な知見であるといえます。

この研究を推進する上で、もっとも困難であった点は、視覚情報を価値情報に変換する過程がどのようなものであるか、神経科学の領域においても、人工知能の領域においても全く確立されていない点でした。私たちは、未知の脳内情報処理過程を解明するために、同様の情報処理を行う人工神経回路を作成し、これをモデルとして用いることで、脳内の神経基盤を同定する、という新規解析パラダイムを考案、確立しました。この成果は、biorxivにて発表され、現在、投稿中です。

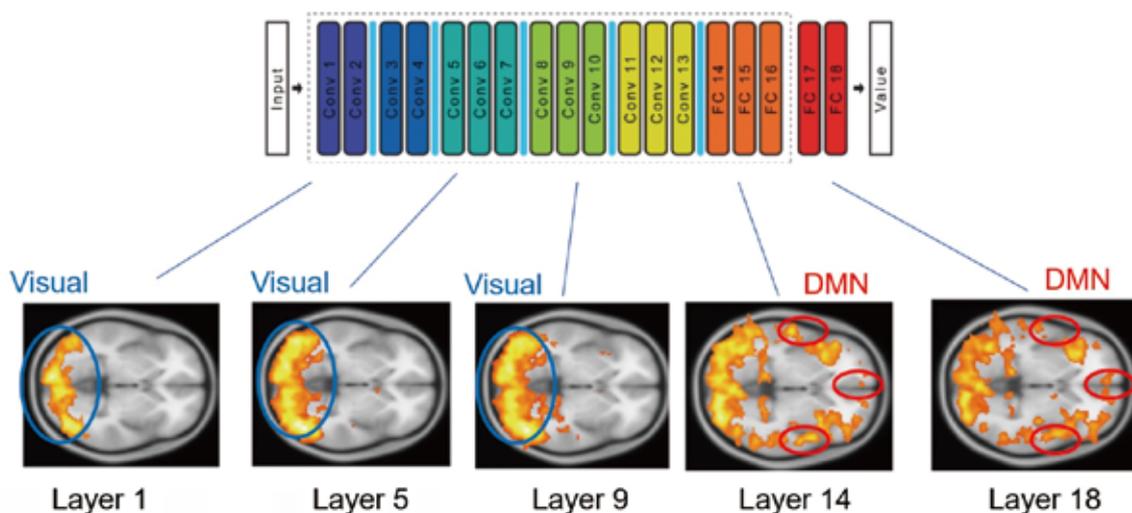


図1 人工知能と脳領域の情報表現の対応。

### 主な研究業績：

- 1) Pham TQ, Yoshimoto T, Niwa H, Takahashi HK, Uchiyama R, Matsui T, Anderson AK, Sadato N, Chikazoe J, Vision-to-value transformations in artificial network and human brains. bioRxiv, 2021  
<https://doi.org/10.1101/2021.03.18.435929>
- 2) Pham TQ, Nishiyama S, Sadato N, Chikazoe J, Distillation of Regional Activity Reveals Hidden Content of Neural Information in Visual Processing, Frontiers in Human Neuroscience, 727, 2021  
<https://doi.org/10.3389/fnhum.2021.777464>
- 3) Chikazoe J, Lee DH, Kriegeskorte N, Anderson AK, Distinct representations of basic taste qualities in human gustatory cortex. Nature communications 10 (1), 1-8. 2019.  
<https://doi.org/10.1038/s41467-019-08857-z>

## 積層独立成分分析の深化と脳科学応用

研究代表者：平山 淳一郎（国立研究開発法人 産業技術総合研究所 人間情報インタラクション研究部門 主任研究員）

多変量データのもつ意味や法則性を理解するためには、データの性質を捉えた表現や特徴量へと適切に変換することが重要です。統計的な基準のもとで変換を定める（学習する）独立成分分析（ICA）や正準相関分析（CCA）といった「教師なし」表現学習の基本的な手法群は、脳・神経科学分野において広く用いられていますが、近年の多様なデータ解析への応用は限られます。また特に非線形で階層性のある表現を得るため、ICAを多段に積み重ねたヒューリスティックな解析法が一部で用いられています（積層ICA）。本研究では、積層ICAを確率的生成モデルの観点から捉え直したSPLICEモデルの拡張を足がかりに、主に脳活動計測データ解析における教師なし表現学習法の新たな応用を探りました。第一に、時間方向の次元縮減を導入した拡張SPLICEを安静時脳波の教師なし非線

形特徴解析に応用し、同時計測した機能的MRI（fMRI）による脳内ネットワーク活動との相関を確認<sup>1)</sup>、脳波ニューロフィードバックへと応用しました（投稿中）。第二に、CCAの多集合拡張を多被験者・多課題fMRIデータベースの解析に新たに適用し、状態（課題）に依らない被験者固有の特性を表現する全脳の機能的結合モードを同定しました<sup>2)</sup>。

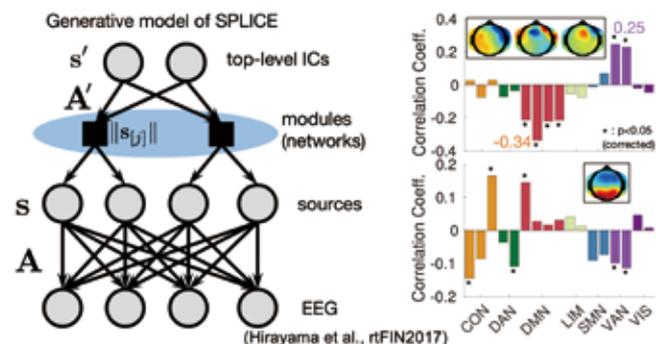


図1 左：SPLICE解析（Hirayama, et al. ICML2017）では階層的な確率的生成モデルの最尤推定により複数の線形特徴抽出層とその間のL2プリーング層からなる非線形特徴抽出器を教師なし学習する。右：拡張SPLICEを安静時脳波データへと適用すると、中間層に同時計測したfMRIによる安静時ネットワークと有意に相関した脳波特徴量が得られた。

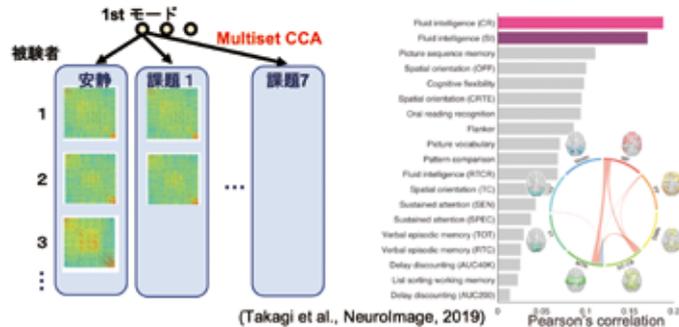


図2 Multiset CCA（Kettering, 1971）をHuman Connectome Projectによる約500被験者のfMRI機能的結合行列に適用し、安静時および7つの課題からなる8状態間で最も再現性が高くなる線形特徴量（モード）を各被験者について求めた。有意に得られた3つのモードは流動的知能など基本的な認知指標との相関を示し、状態によらず被験者の固有特性を反映する基盤的な脳ネットワークへの示唆を与えた。

### 主な研究業績：

- 1) Ogawa T, Moriya H, Hiroe N, Yamada T, Kawanabe M, Hirayama J (2019). Network extraction method using hierarchical ICA-based approach: a simultaneous EEG-fMRI study, Organization of Human Brain Mapping 2019 (OHBM2019)
- 2) Takagi Y, Hirayama J, Tanaka SC (2019). State-unspecific patterns of whole-brain functional connectivity from resting and multiple task states predict stable individual traits. NeuroImage, 201, 116036. <https://doi.org/10.1016/j.neuroimage.2019.116036>

## 単機能の重ね合わせにより新機能を創発するマルチファンクショナル深層学習ネットワーク

研究代表者：柳井 啓司（電気通信大学 教授）

本研究では、単一のニューラルネットワークに複数のタスクのための機能を同時に学習させ、学習した独立の機能の組み合わせ・重ね合わせによって、事前に学習した個別の単一機能とは異なる、新しい機能を実現し、ニューラルネットワークにより人間の脳に近い汎用的な能力を持たせる事が可能であることを実証するための研究を行いました。具体的には画像変換ネットワークを、制御信号の入力によって機能を切り替えることが可能とし、

制御信号の変化させることで同じ画像入力に対して、様々な異なる出力を得られるネットワークを提案しました<sup>1)</sup> (図)。複数の機能を組み合わせも学習可能であることを示しました。また関連して、自己教師あり学習を応用した弱教師あり領域分割<sup>2)</sup>や、画像特徴の意味情報と形状情報を分離する画像と言語に関するクロスモーダル検索<sup>3)</sup>についても研究を行い、それぞれ従来の手法の性能を上回る成果を得ることができました。



図左 提案手法によって1つのネットワークで複数のタスクに対応できるようになりました。

図右 さらに、タスクを混ぜ合わせた混合タスクも処理可能となりました。

### 主な研究業績：

- 1) Mana Takeda, Gibran Benitez-Garcia, and Keiji Yanai (2021). Training of Multiple and Mixed Tasks With A Single Network Using Feature Modulation, Proc. of ICPR Workshop on Deep Learning for Pattern Recognition (DLPR) .  
[https://doi.org/10.1007/978-3-030-68790-8\\_55](https://doi.org/10.1007/978-3-030-68790-8_55)
- 2) Wataru Shimoda and Keiji Yanai (2019). Self-Supervised Difference Detection for Weakly-Supervised Semantic Segmentation, Proc. of IEEE/CVF International Conference on Computer Vision (ICCV).  
<https://doi.org/10.1109/ICCV.2019.00531>
- 3) Yu Sugiyama and Keiji Yanai (2021). Cross-Modal Recipe Embeddings by Disentangling Recipe Contents and Dish Styles, Proc. of ACM Multimedia.  
<https://doi.org/10.1145/3474085.3475422>

## 皮質脳波ビッグデータによる革新的人工知能の開発

研究代表者：柳澤 琢史 (大阪大学高等共創研究院 教授)

ラボHP <https://www.med.osaka-u.ac.jp/pub/nsurg/yanagisawa/>

本研究課題では、自由行動下でのヒトの運動や視聴覚内容および思考内容を大量のライフログデータとして記録しつつ、これと同期して頭蓋内に留置した電極から皮質脳波を計測し皮質脳波ビッグデータを作成しました。各被験者につき10日程度の連続計測を50人以上施行し、のべ500日以上同時計測データを取得しました(図1)。更に、計測されたビッグデータを教師信号として、深層学習等の人工知能技術を用いて、脳信号からヒトの行動・思考・認知内容の全てを包括して予測する脳情報解読を行うことで、脳情報処理の解明を目指しました。視覚情報についてはヒトが知覚する動画の意味内容に

ついてword2vecを用いて意味ベクトルを作成することで、脳活動と対応づけ、皮質脳波から知覚画像の意味を推定できることを明らかにしました<sup>1)</sup>。また、同技術を想起画像推定に応用し、想起した画像内容を画面に提示する技術を開発しました(投稿中)。また、脳活動とAIとの対応関係を変えることで、Neurofeedbackにより運動皮質表象を制御でき、それに伴い幻肢痛を治療できることを明らかにしました<sup>2)</sup>。今後、脳とAIを融合する技術を開発し、その医療応用を進めていきたいと考えています。

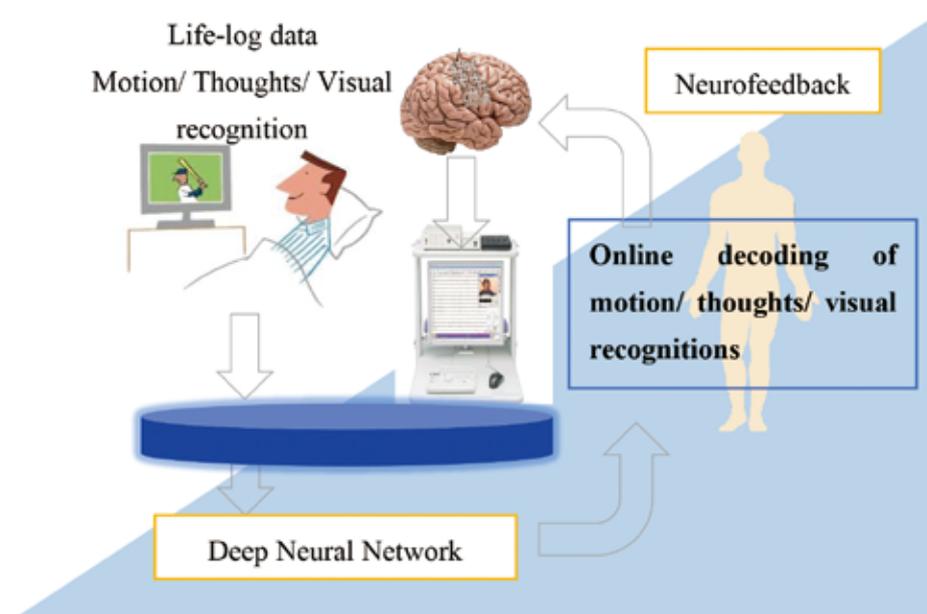


図1 研究の概略図

頭蓋内脳波を計測しながら様々なライフログデータを同時計測し、得られたビッグデータに深層学習を適用することで、高精度な脳情報解読を実現する。さらに、抽出された脳情報をもとにニューロフィードバックを行い、疾患の治療などに応用する。

### 主な研究業績：

- 1) Ryohei Fukuma, Takufumi Yanagisawa, Shinji Nishimoto, Masataka Tanaka, Shota Yamamoto, Satoru Oshino, Yukiyasu Kamitani, Haruhiko Kishima (2018) Decoding visual stimulus in semantic space from electrocorticography signals, IEEE International conference on systems, man, and cybernetics (SMC)
- 2) Takufumi Yanagisawa\*, Ryohei Fukuma, Ben Seymour, Masataka Tanaka, Koichi Hosomi, Okito Yamashita, Haruhiko Kishima, Yukiyasu Kamitani, and Youichi Saitoh (2020) BCI training to move a virtual hand reduces phantom limb pain, Neurology, 95(4):e417-e426, doi: 10.1212/WNL.00000000009858

## 自己シミュレーションとホメオスタシスを基底とする脳のモデル研究

研究代表者：池上 高志（東京大学総合文化研究科 教授）

認知地図という考え方があります。これはアメリカの心理学者であるトールマンによって提唱されたもので、動物は地図のような空間的な表象をもっており、それを用いてナビゲーションしているというものです。オキーフらは、環境のなかのある特定の位置に来たときのみ発火する細胞があることを見出し、これを場所細胞と名付け、これが認知地図を担っていると考えられるようになりました。近年ではこの空間的な認知が他の認知にも利用されていることが見出されており、例えば音の周波数、単純な図形、さらには社会的な上下関係といったものも、同様のシステムで脳内にマッピングされることがわかってきました。

さて、このような空間的な認知はどのように生じるのでしょうか。そのひとつのメカニズム候補として、パス・インテグレーション（経路積分）があります。これによると、動物は自分の動きをベクトル的に足し合わせることで現在の位置を割り出しているとされています。ただし、これでは実空間以外のものに適用できないし、さらに自分の動きの感覚を失われた場合には、空間のマッピングは歪むものの、それでもある程度対応関係が維持されることが報告されていますが、これにも適用できません。

深層学習の分野では変分自己符号化（VAE）や敵対的生成ネットワーク（GAN）といった画像の生成モデルが発展してきました。これらのモデルの特徴として、画像の生成系の入力に、潜在変数という低次元のベクトルを用いることがあります。すなわち、画像という高次元のものを低次元のベクトル空間である、潜在変数空間に圧縮しているのです。これらの出力例を見ると、入力画像の特徴を捉えて自然に低次元空間にマッピングしているように見えます。これから、脳内の認知地図は自分の動きという情報は必ずしも必要ではなく、生成系の学習にとまって自己組織化されるのではないかと考えるに至りました。

そこで、VAE/GANという生成系の深層学習（図1）を、シミュラ的な仮想空間を動き回って取得した一人称視点の画像を用いて予測の学習をおこない、その結果として潜在変数空間

にどのような構造が生じるかを調べることにしました。（図2）これを単純に描画すると環境の地図とは異なるパターンがあらわれており、一見マップはつくられていないようにみえました。しかし、予測した画像の類似度と、潜在変数空間での距離の関係を調べると、これらが相関することを見出しました。すなわち、対象間の近さが、潜在変数空間上の近さと対応している、ということであり、これはマップとなるための重要な要件です。さらに、GANというシステムの働きによって、マップが滑らかとなり、大まかな座標とも言うべきものが生じることもわかってきました。

以上の結果は、近年見出されているさまざまな認知地図について整理する手助けになると考えています。例えば、入力距離を反映しているかどうか、というのが最も単純なマップであり、これは生成系の学習さえすれば実現できます。方向、ないしはおおまかな座標がうまれるためには、マップをなめらかにする働きとしてのGANが必要です。実際の位置との対応を精密に得るためには、今回のシステムでは難しく、やはり運動の情報が必要となるでしょう。認知地図を見出した、という近年の様々な実験結果をこれらの観点で精密化することで、背後のメカニズムへの新たな洞察が得られると期待しています。

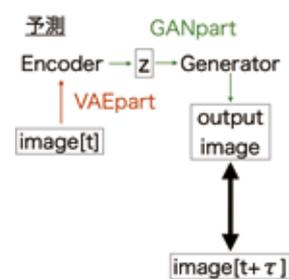


図1 VAE/GANの構成図

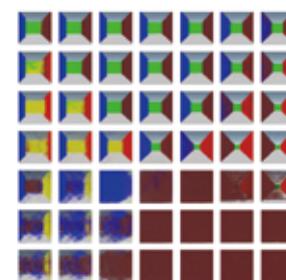


図2 潜在変数空間と出力画像の対応の例。次に予測される壁や廊下の風景を、それを出力する潜在変数空間の場所に対応させて描画している。

### 主な研究業績：

- 1) Kojima H., Ikegami T. (2021). Artificial Cognitive Map System based on Generative Deep Neural Networks. ALIFE 2021: The 2021 Conference on Artificial Life. Online. [https://doi.org/10.1162/isal\\_a\\_00462](https://doi.org/10.1162/isal_a_00462)
- 2) Kojima, H., & Ikegami, T. (2021). Organization of a Latent Space structure in VAE/GAN trained by navigation data. arXiv preprint arXiv:2102.01852.
- 3) Masumori, A., Maruyama, N., & Ikegami, T. (2021). Personogenesis through imitating human behavior in a humanoid robot "alter3". *Frontiers in Robotics and AI*, 165.

## 予測の神経基盤：広域皮質脳波における時空間構造 — 霊長類の脳内で行われる階層的予測符号化 —

研究代表者：小松 三佐子（理化学研究所脳神経科学研究センター 研究員）

予測符号化は大腦新皮質に外界のモデルを獲得する際の普遍的計算原理であると考えられています。しかしながらその神経基盤についてはまだ解明されていません。その原因のひとつとして予測符号化に関わる神経活動を観察する手法が確立されていないことがあげられます。予測符号化は感覚情報が入力されてから数十～数百ミリ秒の単位で起こる動的な情報処理であること、おそらく前頭葉からのフィードバックを含む皮質の広域ネットワークが関係することから、その神経基盤の解明のためには大腦新皮質の活動を高い時空間解像度で広域に観察する必要があります。

本研究では、予測符号化が霊長類の脳内でどのように行われているのかを大域的に明らかにするために、広域皮質脳波電極の開発を行いました。この電極により霊長類大腦皮質広域の情報流を可視化することができるようになりました。また、fMRIと組み合わせることで、霊長類の脳内で行われる階層的予測符号化に関わる脳領域を全脳で同定し、それら領域間の関係を皮質脳波信号から明らかにしました。この研究では、Local-global paradigmという、音の塊の中で生じる「局所的な逸脱」とタスクのコンテキストの中で生じる「大域的な逸脱」との2種類の認知的に異なる逸脱が含まれる聴覚刺激を聞いて

いる動物から脳活動を計測しました(図1)。fMRIの結果「局所の逸脱」では皮質下および低次聴覚野を賦活し、「大域的な逸脱」は高次聴覚野および前頭葉が関わっていました。また、ECoGの活動からは、「局所の逸脱」に対しては低次聴覚野で早い潜時の高周波帯域の活動がみられ、「大域的な逸脱」では高次聴覚野で遅い潜時の高周波帯域の活動、前頭葉では遅い潜時にβ帯域の活動に変化がみられるなど、異なる時間窓でそれぞれの逸脱が処理されていることに加え、「大域的な逸脱」に関わる高次聴覚皮質と前頭葉では役割が異なることを示唆する結果が得られました(図2)。さらにこれらの活動の相関解析などを通し、大域的な逸脱検出に関わる前頭葉の活動が、その次のステップの音に対する聴覚皮質の神経応答に影響を与えていることから、前頭葉の活動が内部モデルの更新に関わっていると考えられました。この研究を通して、霊長類の階層的な予測符号化に関する神経基盤の理解が進んだだけでなく、fMRIで皮質下も含めた脳全体の関連部位を同定し、ECoGでそれらの時間的な関係を明らかにするといった相補的な研究の有効性も示すことができました。このようなアプローチを通して全脳での情報処理様式が明らかになれば、新しいAIアーキテクチャの探求にも繋がると考えています。

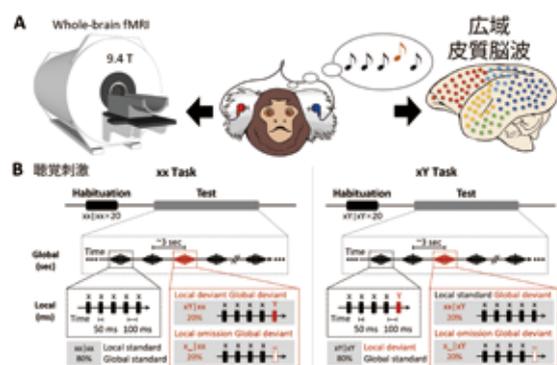


図1 実験デザイン

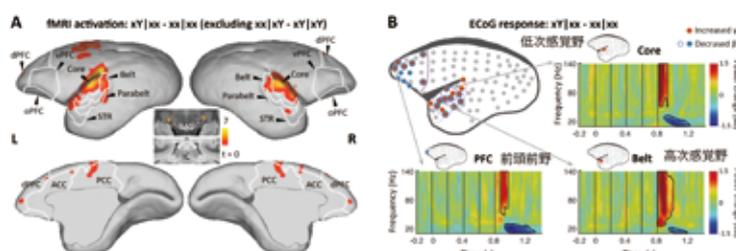


図2 A: 聴覚予測に関わる脳領域 B: 予測の階層性を反映し時空間的に異なる皮質の活動

### 主な研究業績:

- 1) Komatsu M, Ichinohe N (2020). Effects of Ketamine Administration on Auditory Information Processing in the Neocortex of Nonhuman Primates. *Front Psychiatry* 11, 826. doi: 10.3389/fpsy.2020.00826.
- 2) Kaneko T\*, Komatsu M\*, Yamamori T, Ichinohe N, Okano H (2022). Cortical neural dynamics unveil the rhythm of natural visual behavior in marmosets. *Commun Biol* 5 (1), 1-13. \*equal contribution
- 3) Jiang Y\*, Komatsu M\*, Chen Y, Xie R, Zhang K, Xia Y, Gui P, Liang Z, Wang L (2022). Constructing the hierarchy of predictive auditory sequences in the marmoset brain. *eLife*. Accepted. \*equal contribution (この研究は中国科学院のWang博士との共同研究で、本研究領域若手派遣事業の支援を受けて実施しました。)

## 意思決定過程と内部モデルの相互作用

研究代表者：濱口 航介 (京都大学大学院医学系研究科 生体情報科学講座 講師)

私が工学部出身ながら脳科学を志したのは、自ら学習する脳の仕組みを機械に導入すれば、ロボットの「こころ」が作れるのではないかと考えたことが発端でした。人工知能でも脳でも、単純な刺激-応答を超えた知的な行動を行うには、環境の(動的)構造(例:重力や慣性、壁や扉の構造、他のエージェントの行動)を理解し、各々の状態遷移を予測しつつ行動する事が求められます。将来起こりうる状態の組み合わせ爆発を直接に扱う事なく、どのようにして行動を決定していくのか、私はその答えを脳に求めて研究しています。新学術「人工知能と脳科学」では、鼠(マウス)が環境の動的構造を学習し、行動選択に取り入れる仕組みを中心に研究を行いました。これまでに、脳の様々な領域で、環境の構造や知識をベースにした行動の価値(予測的な価値)が神経活動として検出されています。また環境の動的構造を学習する前には、過去の行動-報酬履歴から計算できる行動の価値(回顧的な価値)が行動選択の拠り所です。この予測的および回顧的な価値は、異なる脳領域で計算されている、と考えられています。しかし異なる脳領域で計算された価値が、最終的にどのように統合され、行動を決定しているのか、その仕組みはわかっていませんでした。私達はマウスが課題に習熟するにつれ予測的な行動を行う事、その際、

高次運動野の神経活動が回顧的な価値に加えて予測的な価値を統合している事を、カルシウムイメージングおよび光遺伝学に基づく因果関係の証明を通じて、明らかにしています。新学術「人工知能と脳科学」は、学習や知的行動の原理解明を目指す専門家の集まる場でした。そのような環境に身を置き、共同研究の機会やご指導を頂く事ができたのは、本当に幸運な事でした。この場を借りて、感謝の意を表します。

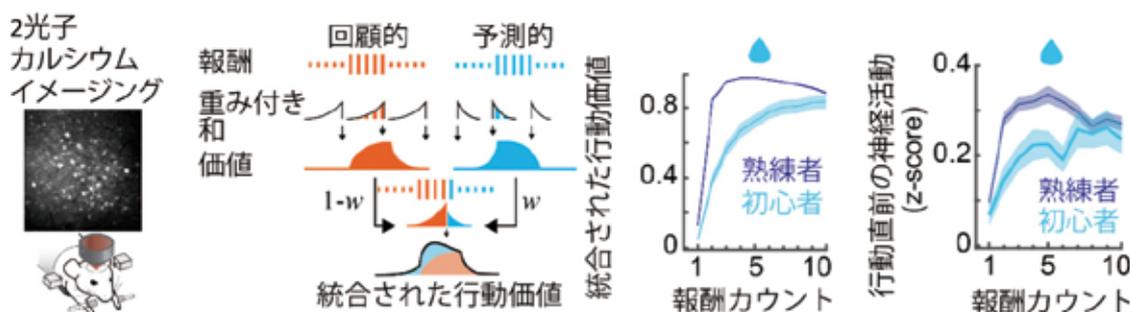


図1 2光子カルシウムイメージング法により同定された予測的価値を表現する前頭皮質領域

### 主な研究業績：

- 1) Hamaguchi K, Takahashi H, Watanabe D (2021). Prospective Action Selection and Inferred Value in Mouse Frontal Cortex, The 44th Annual Meeting of Japanese Neuroscience Society (JNS 2021).
- 2) Nishioka T, Hamaguchi K, Yawata S, Hikida T, Watanabe D (2020). Chemogenetic Suppression of the Subthalamic Nucleus Induces Attentional Deficits and Impulsive Action in a Five-Choice Serial Reaction Time Task in Mice. *Front. Syst. Neurosci.*, 14, 38.
- 3) 濱口航介 (2021). 細胞内膜電位から見た小鳥の歌を紡ぐネットワーク. *日本神経回路学会誌*, 28, 3, 136-143.

## 樹状突起計算を統合する新しい回路学習パラダイムの創出

研究代表者：深井 朋樹（沖縄科学技術大学院大学 教授）

前期は「大脳皮質局所回路に学ぶ新しいアーキテクチャと学習モデルの構築」、後期は「樹状突起計算を統合する新しい回路学習パラダイムの創出」という課題を掲げて研究しました。まず初めに、過去にラットの内側前頭野（MFC）から記録した多電極記録データを用いて、音刺激依存に行動選択を学習するメカニズムをモデル化しました<sup>1)</sup>。特に過去の研究であまり扱われて来なかった、意思決定の個体差と大脳皮質の神経ダイナミクスの関係性を問いました。高低二種類の学習された刺激音と、MFC神経活動の記録時に挿入された新奇音に対するラットの行動選択確率を観察すると、新奇音に対する行動選択はラット毎に大きな違いを見せます（図1）。意思決定は神経集団活動の時間発展（神経軌道）により支配されることが知られていますが、「心理測定曲線」に個体差が生じるメカニズムを明らかにするため、神経回路モデルを構築してラット同様に訓練しました。その結果、「心理測定曲線」はMFCの神経軌道が外的擾乱の影響を受けやすい場合は滑らかに変化し、影響を受けにくい場合は強いバイアスやランダム性を示すことがわかりました（図1）。少し逆説的に聞こえますが、一見賢く見えるラットの脳の方がノイズに弱い事もあり得ることを示唆しています。

後期は樹状突起の情報処理に研究の軸足を移しました。樹状突起は領野間の情報の流れやシナプス可塑性の制御に関わる大切な計算要素であり、最近、ディープリンクとの関連性なども調べられています。樹状突起は生物学的には複雑な「マシン」ですが、この新学術領域の目標はAIと神経科学の両方に跨りますので、生物学的詳細に拘わらず大胆なアイデアを試しました。現実世界の情報を扱うには異なるスケールに亘る階層構造（チャンク構造）を読み解かねばなりません。そ

こで、この能力を実現する神経メカニズムを研究しました。

まず細胞体と樹状突起から成る2コンパートメントモデルを構築して、単一神経細胞によるチャンクの抽出を実現しました<sup>2)</sup>。鍵となる原理は「樹状突起と細胞体間に於ける情報損失の最小化」です。この学習原理が入力データに対する尤度最大化と異なる点は、樹状突起が学習すべきデータが“逆伝搬スパイク”を介して、細胞体の発火活動により生成される点です。細胞体が樹状突起による学習の教師を務めることで、ニューロン全体で教師なし学習を実現します（図2）。興味深いのは、同一の回路モデル（学習原理）を用いて、Blind Source Separation（信号音源分離）も実行可能なことです。提案した単一細胞モデルを神経回路の学習理論に組み込み、学習能力を著しく向上させることも成功しました。

最後に古典的な連想記憶モデルを発展させて、空間的階層構造の学習についても研究を行いました。このモデルでは、相関構造を持つアトラクタ状態の生成と、2つの抑制性神経回路による相関距離の調節により階層構造の柔軟な想起を実現します。このような想起により、例えば迷路探索においてはボトルネック（=サブゴール）の発見が可能になります（図3）。樹状突起のモデルと併せて、時空間情報の階層構造の学習に関するレビュー論文を発表しました<sup>3)</sup>。今後、ここで得られた学習メカニズムを予測符号化モデルなどに統合して、脳が適切を“統計変数”を発見し、外界や自己の統計モデルを構築する過程をモデル化したいと考えています。また空間学習モデルを言語解析に発展させる可能性を検討しています。

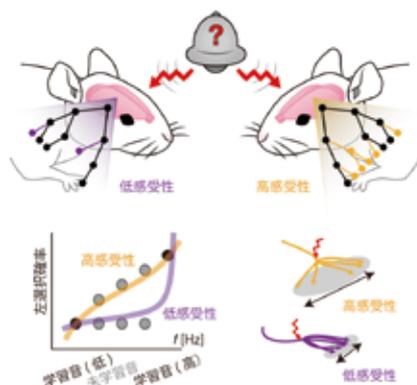


図1 意思決定行動の個体差と神経活動の個体差の関連。左下は心理測定曲線。

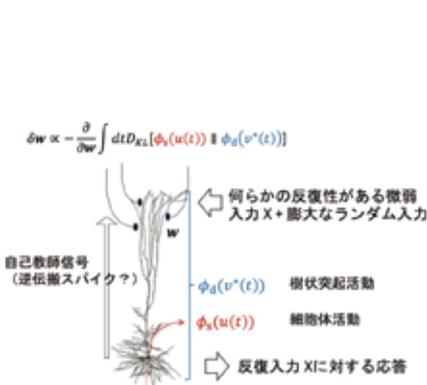


図2 単一神経細胞モデルによる時系列構造の自己教師学習モデル

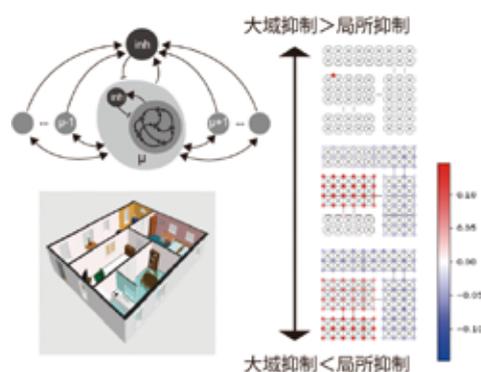


図3 相関型アトラクタを用いた空間的階層構造の記憶神経回路モデル

### 主な研究業績：

- 1) Kurikawa, T., Haga, T., Handa, T., Harukuni, R. & Fukai, T (2018). Neuronal stability in medial frontal cortex sets individual variability in decision-making. *Nature Neuroscience*, 21, 1764-1773.
- 2) Asabuki, T., Fukai, T (2020). Somato-dendritic consistency check for temporal feature segmentation. *Nature Communications*, 11, 1554, doi:10.1038/s41467-020-15367-w.
- 3) Fukai T., Asabuki, T., Haga, T (2021). Neural mechanisms for learning hierarchical structures of information. *Current Opinion in Neurobiology*, 70, 145-153.

## Learning perceptual representations in biological and artificial neural networks

研究代表者: Andrea Benucci (RIKEN Center for Brain Science, Team Leader)

### Stable Visual perception as classification invariance in convolutional neural networks

Our ability to perceive a stable visual world in the presence of continuous movements of the body, head, and eyes has puzzled researchers in the neuroscience field for a long time. Numerous theoretical studies have proposed biologically plausible computational processes to convert location information—"where" type of signals—across multiple frames of reference, for instance, from retinocentric to craniocentric coordinates. However, how coordinate transformations can then be used by the hierarchy of cortical visual areas to produce stable perception remains largely unknown. We reformulated this problem in the context of hierarchical convolutional neural networks (CNN)—whose conception has been inspired by the hierarchical signal processing of the mammalian visual system—and examined perceptual stability as an optimization

process that identifies image-defining features for accurate image classification in the presence of movements. Movement signals, multiplexed with visual inputs along overlapping convolutional layers, aided movement-related classification invariance by making the classification faster to learn and more robust relative to input noise. Classification invariance was reflected in activity manifolds associated with image categories emerging in late CNN layers and with network units acquiring movement-associated activity modulations as observed experimentally during saccadic eye movements. Our findings provide a computational framework that unifies a multitude of biological observations on perceptual stability under optimality principles for image classification in artificial neural networks (Benucci A., PLOS Comp. Biology, 2022).

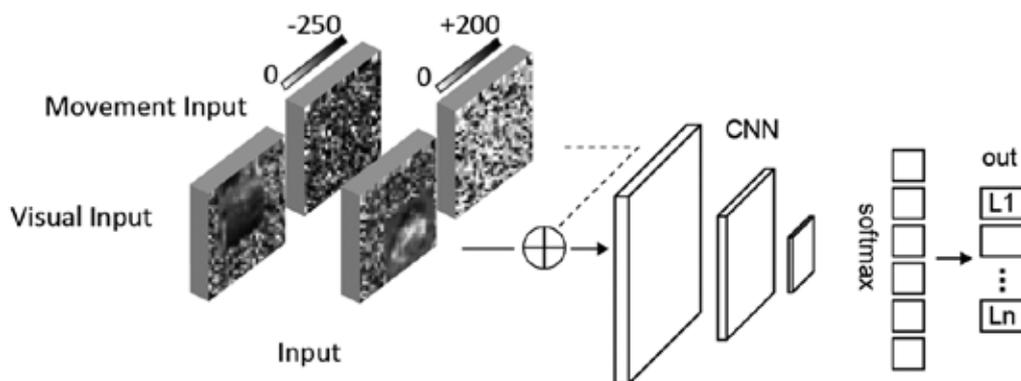


Figure 1. Neural network architecture. movement signals are provided as inputs to a convolutional neural network together with CIFAR-10 images, xy-shifted to mimic eye movements. The amplitude of the movement input is proportional to the image shift.

#### 主な研究業績:

- 1) Orlandi J.G., Abdolrahmani M, Aoki R, Lyamzin D.R., Benucci A (2021). Distributed context-dependent choice information in mouse dorsal-parietal cortex, bioRxiv, DOI:10.21203/rs.3.rs-288103/v1
- 2) Lyamzin D.R, Aoki R, Abdolrahmani M, \*Benucci A (2021). Probabilistic discrimination of relative stimulus features in mice, PNAS 118 (30) e2103952118; <https://doi.org/10.1073/pnas.2103952118>
- 3) Abdolrahmani M, Lyamzin, D.R., Aoki R, \*Benucci A (2021). Attention Decorrelates Sensory and Motor Signals in the Mouse Visual Cortex, Cell Reports, Volume 36, Issue 2, <https://www.sciencedirect.com/science/article/pii/S2211124721007750>

## セロトニン神経系の光操作・観察によるモデルベース的意思決定の脳内機構解明

研究代表者：大村 優 (北海道大学医学研究院 講師)

直接経験だけに基づいて行動しては、刻々と変化する複雑な環境に適応することが難しくなります。既知の限られた情報から推測することは環境適応に重要な知的能力だと考えられます。例えば、A→B→Cという情報を得た後に、B→Dという情報を別途得たとします。この場合、通常ヒトはA→B→Dという一連の情報を直接得なくても、そうだろうと推論します。本研究では、「推論過程に脳内セロトニンが関与する」という仮説を、マウスで検証しました。マウスには、①「光る穴に鼻を入れると餌が貰える」ことをまず学習させます。そして別の日に、単に餌を食べさせた後に、気持ちが悪くなる薬(リチウム)をマウスに投与します。そうすると、マウスは②「この餌を食べたら気持ちが悪くなる」と関連付けます。以上の状況を別々に経験させた後に、再び光る穴に鼻を入れられる状況にマウスを置きます。するとマウスは①と②の経験から推論し、穴に鼻を入れる回数は減ります。

さらに、この推論しているであろうタイミングで背側縫線核のセロトニン神経活動を抑制すると、マウスが推論をしなくなりました。このように、直接経験していないことも脳内でシミュレーションして推論する、という生物の生存に重要な知的機能に一部のセロトニン神経が関与すると考えられます。

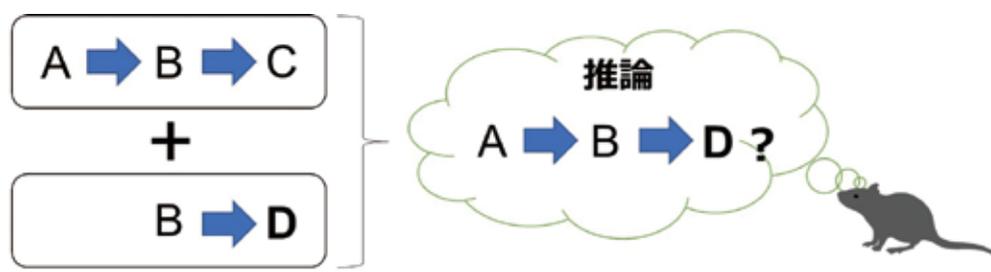


図 背側縫線核のセロトニン神経活動が抑えられると推論をあまりしなくなり、単に直接の経験に基づいた行動を見せる。

### 主な研究業績：

- 1) Ohmura Y, Iwami K, Chowdhury S, Sasamori H, Sugiura C, Bouchekioua Y, Nishitani N, Yamanaka A, Yoshioka M. Disruption of model-based decision making by silencing of serotonin neurons in the dorsal raphe nucleus. *Current Biology*, 31, 1-9, 2021.

## 多次元の状態・行動空間における意思決定と大脳基底核の情報表現

研究代表者：鮫島 和行（玉川大学脳科学研究所 教授）

私たちの研究では、「対象の選択」という一見聞き慣れない言葉の神経表現を実験的に示しました。私たちは日常的に意思決定を行なっています。特に価値に基づく意思決定は、過去の経験に基づいて未来の選択の結果を報酬予測（価値）として予測し、その比較によって合理的に意思決定を行うと説明されます。これまでの神経生理学的な研究では、具体的な運動（左右の眼球運動や、手の動き）の中から1つを選んで行動し、その後得られる報酬の確率や量を実験的に操作することによって、その時の脳内の表現と行動選択との対応関係を研究するものでした。大脳基底核はこの様な実験から行動の価値情報を持つことが知られています。しかし一方で、私たちはより抽象的な対象を選択することもできます。昼食を中華にするのかイタリアンにするのか、具体的なお店や位置を把握し行動に移す前に決めることができます。運動を行う前に、より抽象的なカテゴリーに対して意思決定を行い、その計画目標に従って状況に合わせて柔軟な行動を行えるのです。将来の目標を定め行動を計画する脳内機構は、これまで前頭前野を含む大脳皮質が重要な役割を果たすことがfMRIなどのヒトの研究で示されてきました。しかし、この前頭前野からの神経投射を受ける大脳基底核の神経回路がどのような役割を果たすのかは未知のままでした。そこで、私たちは「対象の選択」と「運動の選択」が時間的に分離するような課題を開発し、それを動物が実行している際の大脳基底核、特に線条体と呼ばれる大脳皮質からの入力部位の単一神経活動を記録し、その神経活動を解析して、大脳基底核が対象の選択に関わるのかどうかを検討しました。

画面上に2種類の図形を提示し、その1つを選択させる、という課題を

動物に行かせます。ただし、選択をする方法は一度選択肢が消えたあと、一つずつ提示された時に反応する、という方法を用いました（図1）。さらに、反応の後に得られる報酬量は図形の形、または、色のどちらか一方に関連づけられます（そしてそれらは一定期間で切り替えられます）。つまり動物は選択肢の図形で、どの形（または色）が一番お得か（報酬がたくさん得られるのか）を、経験を通じて学習しながら選択することになります。動物が行動を起こす前において、線条体から電気生理学的方法で単一の神経細胞の発火活動を記録したところ、細胞ごとに特定の「選択した色」や「選択した形」に強く反応することがわかりました。これは具体的な運動を選択する前に、抽象的な「選択した対象」の情報表現が大脳基底核にあることを示しています。従来の研究と同様に、線条体には「運動の選択」に関わる神経細胞も見つかりましたが、そのニューロンたちと「対象の選択」に関わるニューロンたちは別々に存在していることもわかりました。このことは、将来しなければならない行動を計画的に実行するための抽象的な選択に関わる部分と、具体的に今選択しなければならない運動の選択に関わる部分が、大脳基底核の中に独立して存在することを示しています。

大脳皮質と同様に大脳基底核の線条体の中にもこのような階層性が存在し、それぞれが大脳皮質と連絡を取りながら我々の柔軟な意思決定が実現されているのではないかと考えています。さらに、この細胞活動の解析を進め、これらの抽象的な対象への情報表現がどのように学習によって獲得されていくのか、そのメカニズムを現在調べています。また、このような神経情報表現の研究は、機械に階層的な意思決定を実現するための鍵となる脳メカニズム解明への糸口となる成果です。

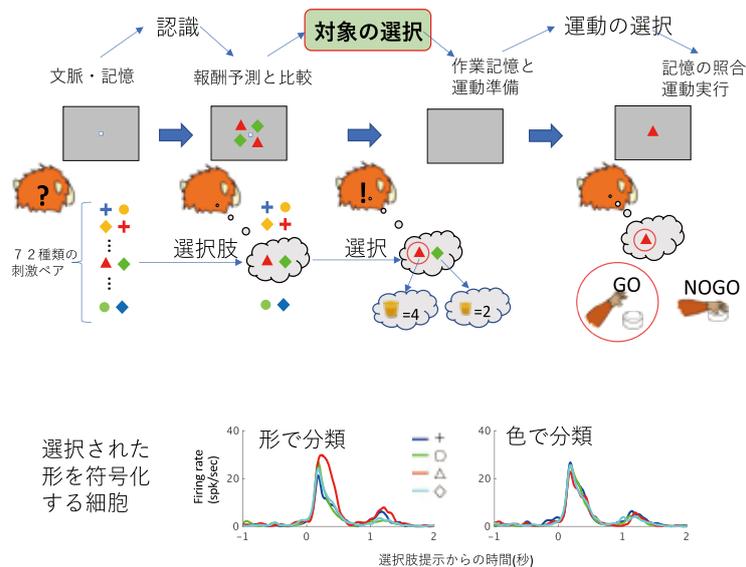


図1 対象の選択と運動の選択を時間的に分離する課題。(上段) 72種類の刺激ペアの中から1つのペアが画面上に提示され、2つの刺激の報酬予測（この場合赤三角の報酬が緑ダイヤモンドの報酬より大きいので、その対象を選択する。その後赤三角が再び提示されたタイミングで運動を行なって、報酬を得る課題。(下段) 選択肢が提示された時のある神経細胞の活動。三角が選ばれる時のみ強い発火が見られる。色ごとに分類しても活動に違いはない。

### 主な研究業績：

- 1) Nonomura, S. & Samejima, K. (2019). Neuronal representation of object choice in the striatum of the monkey. *Frontiers in Neuroscience*, 13, 1283. <https://www.frontiersin.org/articles/10.3389/fnins.2019.01283/full>

## 視覚心理に基づくテクスチャ特徴表現と深層特徴表現のマッピング

研究代表者：庄野 逸（電気通信大学 大学院情報理工学研究科 教授）

### テクスチャ画像を用いた心理物理量と畳み込みニューラルネットワーク表現の関連

本研究では人の視覚心理物理量と深層畳み込みニューラルネットワーク (CNN) の関係を探るべく、心理視覚的な特徴量である Portilla-Simoncelli 統計量 (PSS) を中心にテクスチャ画像記述に関する特徴量研究を行いました。ここでは自然画像のみを学習させた VGG と呼ばれる CNN に対してテクスチャ画像を提示して得られる中間層の表現を計算します。この CNN の特徴表現に対して、テクスチャ画像から得られる PSS 統計量をスパース回帰させ、その重みを調べることにより、PSS が CNN の表現を、どの程度説明可能性のあるのかを測るとい

手法を提案しました (図1参照)。この結果として、VGG 中間層部分では、自然画像のみで学習させたものであるにも関わらず、サル高次視覚野 (V4 野) と類似した表現が発現していることがわかりました。その一方で VGG では、画像中の物体の影などの成分を強調するための変量がほとんど発現しないという特有の性質を持つことも明らかになり、脳の情報表現との違いに関しても明らかになってきました<sup>1)</sup>。また PSS を用いたスピニアウト的な研究として、磁気薄膜画像の計測に PSS を適用することで、物性パラメータの推定などが可能であることを示すことができました<sup>2)</sup>。

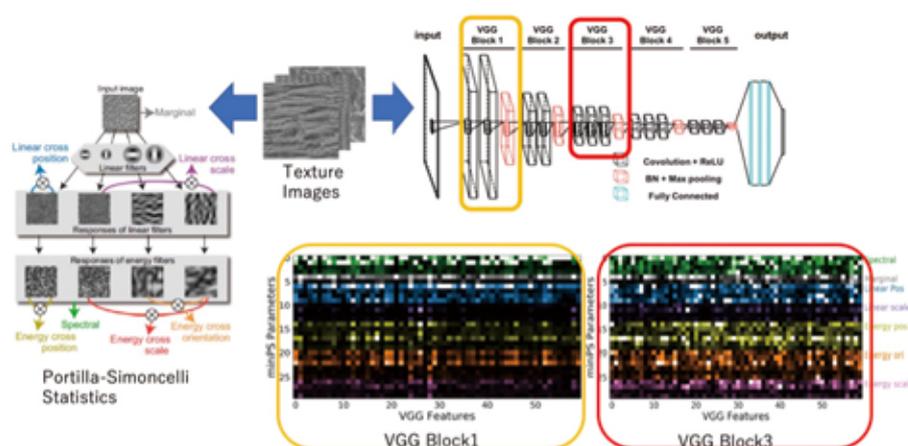


図1 実験概略図。テクスチャ画像の PSS 特徴量と VGG の表現の関連を探る

#### 主な研究業績：

- 1) Hamano Y., Shouno H. (2020), Analysis of Texture Representation in Convolution Neural Network Using Wavelet Based Joint Statistics, Lecture Notes in Computer Science, Proc. of ICONIP' 20, Bangkok, Thai, 12532(1), pp. 126-136  
[https://doi.org/10.1007/978-3-030-63830-6\\_11](https://doi.org/10.1007/978-3-030-63830-6_11)
- 2) Murakami R., Mizumaki M., Hamano Y., Akai I., Shouno H. (2021) Texture Analysis of Magnetic Domain Images Using Statistics Based on Human Visual Perception, Journal of the Physical Society of Japan, 90, 44705,  
<https://doi.org/10.7566/JPSJ.90.044705>

## 高次視覚野の計算理論に基づくワンショット学習モデル

研究代表者：細谷 晴夫（脳情報通信総合研究所 主任研究員）

ワンショット学習は、脳の最も驚くべき汎化能力の一つです。例えば、顔画像を一例だけ覚えておくだけで、どの方向を向いた顔を見せられても認識できます。ワンショット学習は、人工知能分野でも時折取り上げられる難問ですが、決定的なアルゴリズムはまだありません。本研究では、神経科学における高次視覚野の理論的・実験的知見から着想した深層生成学習モデル、Group-based Variational Autoencoder (GVAE) および、その拡張モデルであるCategorical Invariant Generative Model (CIGMO: 右図) を構築し、ワンショットクラス別性能において既存モデルを凌駕することを示しました。

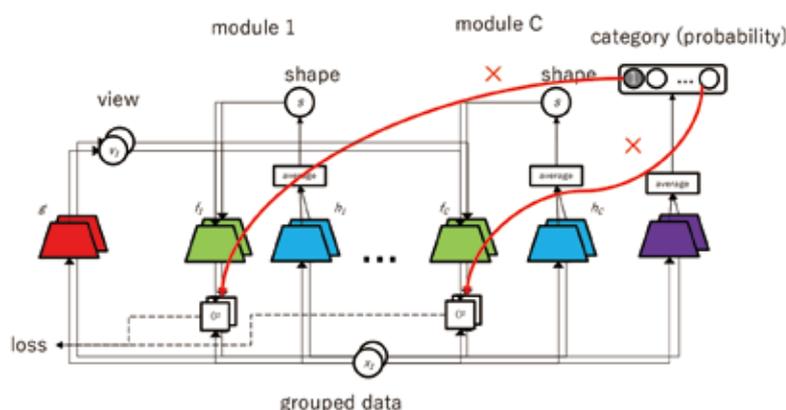


図 CIGMOアーキテクチャ。グループ化された画像データからカテゴリを推定（右端ネットワーク）すると、そのカテゴリに対応した形状表現（中央）が、ビュー表現（左端）と分離するように動作・学習する。

### 主な研究業績：

- 1) Hosoya, H. CIGMO: Learning categorical invariant deep generative models from grouped data. In ICLR Workshop on Weakly Supervised Learning, 2021.  
<https://weasul.github.io/papers/4.pdf>
- 2) Hosoya, H. Group-based learning of disentangled representations with generalizability for novel contents. The International Joint Conference on Artificial Intelligence (IJCAI), Aug, 2019.  
<https://doi.org/10.24963/ijcai.2019/348>
- 3) Raman, R. & Hosoya, H. Convolutional neural networks explain tuning properties of anterior, but not middle, face-processing areas in macaque inferotemporal cortex. Commun. Biol. 3, 1-14 (2020).  
<https://doi.org/10.1038/s42003-020-0945-x>

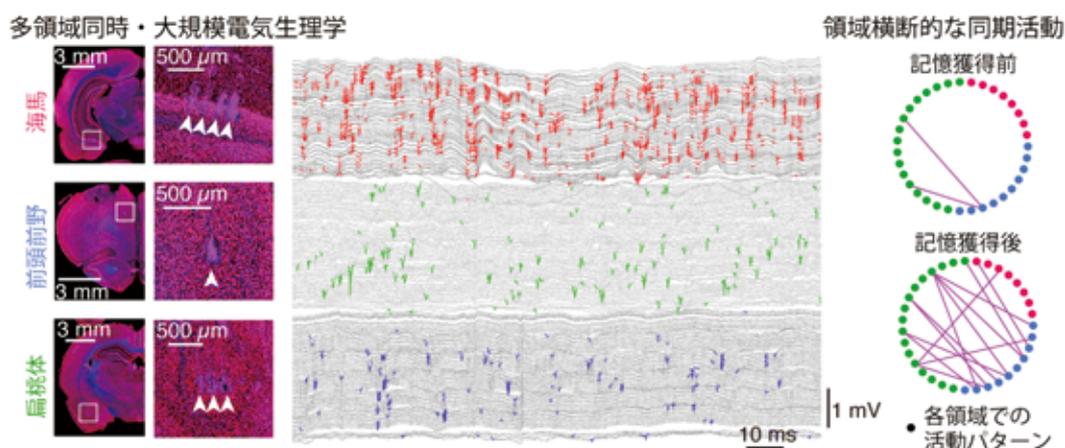
## 超大規模電気生理学を用いた睡眠中のオフライン学習アルゴリズムの解析

研究代表者：宮脇 寛行 (大阪市立大学 大学院医学研究科 助教)

脳にとって睡眠は外界からの感覚入力から切り離されたオフラインの状態です。しかし脳は睡眠中も活動し続けており、覚醒時に見られた神経活動パターンが散発的に再現されていることが知られています。このような活動の「リプレイ」は海馬をはじめとする様々な脳領域で報告されており、覚醒時に獲得した記憶を定着させる「オフライン学習」に関わっていると考えられています。しかし、様々な脳領域でおこるリプレイがどのように記憶を定着させているのかは不明でした。

そこで本研究では、海馬、扁桃体、大脳皮質の3つの領域から多数の神経細胞の活動を一齐に記録し、記憶の獲得によって活動パターンの間に見られる関係がどのように変化するかを解析しました。その結果、記憶の獲得後の睡眠中には複数の脳領域で同期したリプレイが生じることが明らかになりました。さらに興味深いことに、海馬と大脳皮質の活動パターン

は記憶の獲得中には弱い同期しか示さないのに対し、その後の睡眠中に強く同期するようになることを発見しました。また、このような同期活動は記憶の想起の際にも観察されました。これらの結果は、睡眠中のリプレイによって脳の領域横断的なネットワーク結合が強化されることが、オフライン学習に重要な役割を果たしている可能性を示唆しています。



図左 多領域同時記録部位の例。

図中央 同時記録された局所電場電位の例。単一細胞の活動に由来する波形を強調表示している。

図右 記憶の獲得前（上）および獲得後（下）の睡眠中に見られる同期パターン。記憶の獲得にともない、脳領域横断的な同期活動が生じることが明らかとなった。

### 主な研究業績：

- 1) Miyawaki, H., & Mizuseki K. (2019) Evolution of fear memory representation in local and global circuits. (The 49th Annual Meeting of the Society for Neuroscience)
- 2) Miyawaki, H., & Mizuseki K. (2020) Dynamics of inter-regional cellular interactions during slow-wave sleep. (The 43rd Annual Meeting of the Japan Neuroscience Society)
- 3) Miyawaki, H., & Mizuseki K. (2022) De novo inter-regional coactivations of preconfigured local ensembles support memory. Nature Communications, in press.

## 非線形ダイナミカル表現学習法の開発による脳の理解と予測

研究代表者：森岡 博史 (理化学研究所 革新知能統合研究センター 特別研究員)

このプロジェクトでは主に、未知の非線形ダイナミクスに潜む潜在成分をデータ駆動的に推定する手法を提案しました。提案法は、近年我々が開発した深層学習に基づく非線形独立成分分析 (ICA) を、内部に再帰的な結合を持つダイナミカルモデルへと理論的に拡張したものです。一般的に、このような非線形表現学習は極めて不定性の高い問題であり、一般化非線形ダイナミクスにおいて同定性を保証できる推定法はこれまで存在していませんでした。本研究においては、潜在成分が独立同分布でないことを仮定することにより、未知の非線形ダイナミクス (特別な仮定は不要) から得られた観測データからその同定が可能であることを示し、さらに深層学習に基づく実用的な推定法を提案することで注目を浴びました。提案法の有効性を示すため、外部刺激 (視覚・聴覚) 下におけるヒトの脳活動計測データ (MEG) に適用した結果、脳ダイナミクスが実際に外部刺激

に依存する形で駆動されていることが示されました。また、それらの潜在成分の被験者間での類似度を評価した結果、提案法による潜在成分が他手法よりも高い値を持つことが示されました。提案法は非線形ダイナミクスを解析する上で極めて汎用性の高い手法であり、脳に限らず様々なダイナミクスに適用可能であるため、今後さらなる発展が期待できます。

提案法を実装したコードはGitHubにて公開しているので (研究業績参照)、興味がありましたらぜひお試しください。また、要望やフィードバックなどありましたらいつでもご連絡ください。

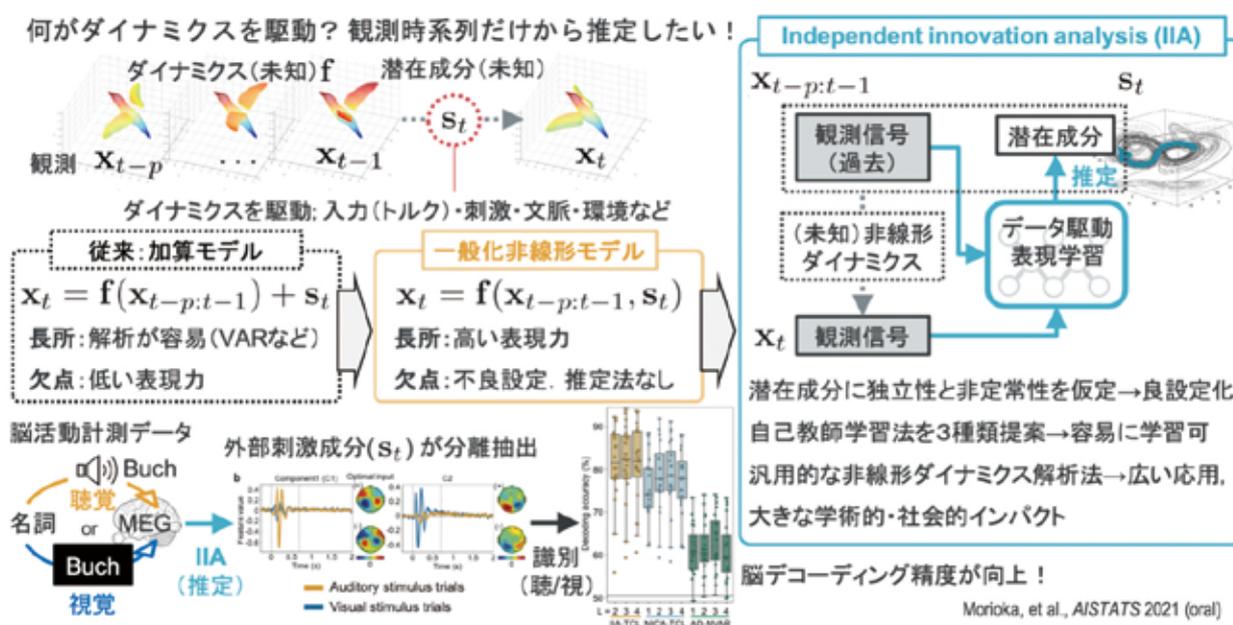


図 非線形ダイナミカル表現学習法の開発による脳の理解と予測

### 主な研究業績:

- 1) Morioka H, Hälvä H, Hyvärinen A (2021). Independent Innovation Analysis for Nonlinear Vector Autoregressive Process. International Conference on Artificial Intelligence and Statistics (AISTATS). <https://virtual.aistats.org/virtual/2021/poster/1692> (code: <https://github.com/hmorioka/IIA>)

## 潜在的運動における学習適応メカニズムの解明と計算モデル構築の研究

研究代表者：五味 裕章 (NTT CS研 人間情報研究部 感覚運動研究グループ リーダ・上席特別研究員)

「思わず体が反応した」というのは、スポーツばかりでなく、日常生活の中でも時折体験することでしょう。例えば、バスの中で体がよろけたときに、手すりを掴むように素早く手を伸ばす動作は、意思決定に基づく随意的動作のみでは説明が付きません。これが可能となるのは、意識的な運動を行う情報処理に加え、複数の無意識的な感覚運動情報処理が並列的に階層的に行われているためです<sup>1)</sup>。次世代ロボット開発などにも役立つことが期待される、隠された脳の情報処理の仕組みを明らかにしていくのが我々の研究の狙いです。ここでは、この研究プロジェクトで行ったいくつかの研究についてご紹介します。

我々は以前の研究で、手到達運動中に視野の動きを与えると、短潜時で手が視野の動きの方向に動くMFRという現象を調べてきましたが、この応答が、どのような計算メカニズムで発生するのかについては論争になっていました。仮説Aは、「視覚運動が到達指標の脳内表象の位置ずれを引き起こし、その疑似的位置ずれに対する修正動作が無意識で生ずる」というものでした。この説は知覚研究でMIPS (motion induced position shift) として知られた錯覚現象によって支持されていました。仮説Bは「視覚運動により誤った身体運動推定が行われ、それに対して手の到達運動を修正する動作が無意識に生ずる」というものでした。これらは長年論争になっていましたが、我々の実験は、1: 視覚運動によって生ずる「指標位置の疑似変化」に対する運動修正成分では、視覚運動によって生ずる腕応答の初期成分が説明できないこと (図1A)<sup>2)</sup>、2: 姿勢や視覚的外界の不安定な環境文脈によるMFRのモジュレーションは、視覚運動からベイズ推定される姿勢の動きに対する補償動作としてモデル化できること (図1B) (投稿中)、を示すことに成功し、仮説Bが支持さ

れました。この結果は、脳が、腕運動とは一見関係の薄い「視覚的外界情報や姿勢の安定性」を考慮し、腕運動の実行中に動作修正に結び付けていることを意味しており、脳の優れた情報処理能力の一つであると言えます。

また別の研究<sup>3)</sup>では、ヒトの歩行における歩行速度調節メカニズムを探りました。従来、歩行速度の調節には、映像のコントラスト時間周波数が関係すると考えられていました。我々の研究では、人工視覚環境 (図2) を使って目に映る壁の情報 (奥行き、時空間周波数) や単眼・両眼の情報を操作し、ヒトが視覚の奥行き情報まで考慮した歩行速度の推定を行い、無意識に歩行速度調整を行っていることを明らかにしました。これにより、無意識な視覚—運動制御情報処理であっても、高度な視空間情報処理をおこなって歩行が制御されていることが分かってきました。

さて、視覚情報は、高度な運動制御に直接的に使われることに加え、体性感覚による運動制御にも間接的に影響を与えます。Newsletter vol.9でご紹介したように、手首の運動中に外力を与えて生ずる短潜時の運動応答 (伸長反射) は、手首の運動中の位置カーソルの実際の動きとの差やカーソル消失で変化することから、視覚情報に依存した手の状態表現の曖昧性に応じた体性感覚—運動系の調整が行われていることがわかってきました。このように、視覚情報は、顕在的な外界の幾何学的世界を再構成するというのではなく、外界とのインタラクションを行う運動のための様々な情報処理に潜在的に使われています。今後も我々は、様々な階層レベルで行われる隠された感覚運動制御のための脳の情報処理を明らかにしていくとともに、構成的な検証も行い、人間の感覚運動制御の仕組みの理解を深めていきたいと考えています。

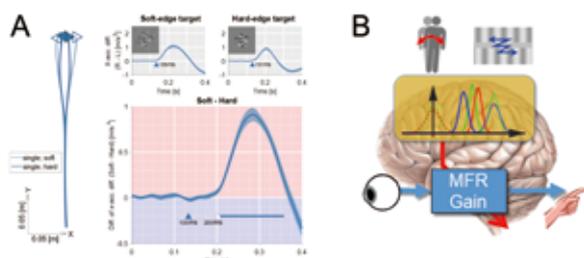


図1 視覚運動の腕応答への影響 A: 輪郭のぼやけたターゲットと輪郭のはっきりしたターゲットの内部運動刺激に対する腕運動中の視覚性運動応答 (上図)と、それらの差分から求めた運動—位置インタラクションに対する成分 (下図) B: MFRの生成プロセスを視覚運動からベイズ推定される姿勢の動きに対する補償動作としてモデル化すると、そのゲイン変化がうまく説明できる

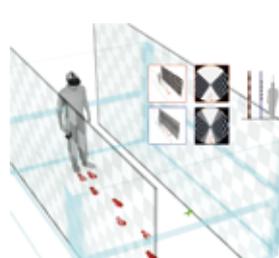


図2 HMDによりバーチャル空間を呈示した際の歩行計測実験模様のパターン・壁の近さ・見え方を操作した際の視覚速度変化に対する歩行速度の調節を観察した



図3 UCL-ICN, Gatsby, NTT, AIBS Joint Workshop: Analysis and Synthesis for Human/Artificial Cognition and Behavior (2018.10 OISTにて)

### 主な研究業績:

- 1) Macpherson T, Matsumoto M, Gomi H, Morimoto J, Uchibe E, Hikida T (2021) Parallel and hierarchical neural mechanisms for adaptive and predictive behavioral control. *Neural Networks* 144:507–521.
- 2) Ueda, H., Abekawa, N., Ito, S., & Gomi, H. (2019). Distinct temporal developments of visual motion and position representations for multi-stream visuomotor coordination. *Scientific Reports*, 9(1), 12104.
- 3) Takamuku S, Gomi H (2021) Vision-based speedometer regulates human walking. *iScience*:103390.

## 報酬/目的指向行動の神経回路機構

研究代表者：足田 貴俊 (大阪大学蛋白質研究所 高次脳機能研究室 教授)

報酬/目的指向行動を適切にかつ柔軟性を持って遂行する機構の解明は脳科学と人工知能に共通する課題です。大脳基底核は報酬/目的指向行動に必須な脳部位であり、パーキンソン病、薬物依存症、PTSD、うつ病、統合失調症といった多くの精神神経疾患で障害されています。研究代表者は、これまでに大脳基底核の直接路と間接路の2つの主要な神経経路について、特異的な可逆的神経伝達阻止法を用いて、マウスにおいて、直接路は報酬学習、間接路は忌避学習にそれぞれ重要な役割を担っていることを示してきました。しかしながら、この並列回路の制御機構については未解明な点が残されていました。

本研究において、大脳基底核の直接路と間接路のそれぞれに特異的な可逆的神経伝達阻止法をマウスに適用することによって、報酬/目的指向行動における大脳基底核神経機構を調べてきました。タッチスクリーン認知行動測定装置を用いて、側坐核の直接路が報酬/目的指向行動の特にサイン追跡行動に関与することを示しました (Frontiers in Neuroscience, 2018) また、マウス集団行動を観察するインテリケージ認知行動測定装置を用いて、行動の柔軟性に側坐核の間接路と間接路に局在するドーパミンD2受容体が重要であることを示しました (Learning & Memory, 2016)。これらの一連の研究から、直接路と間接路の並列回路が適切かつ柔軟な行動に重要であることを明らかにしました。また、側坐核の直接路と間接路は大脳皮質内側頭前野に

おいて、それぞれ回路特異的な遺伝子発現制御を行っていることを示し、階層回路制御の分子機構の一端を明らかにしました (Scientific Reports, 2020)。本研究で明らかとなった報酬/目的指向行動の神経機構をもとに、A02領域において、柔軟な行動に大脳基底核の並列回路および大脳皮質を含めた階層回路が重要であることを概説し、運動制御における人工知能・ロボット制御への応用を議論しました (Neural Networks, 2021)。今後の脳科学に基づく人工知能の開発につなげていきます。

精神疾患では報酬/目的指向行動を適切にかつ柔軟性を持って遂行することが困難になります。本研究において、精神疾患モデルマウスでは大脳基底核の間接路に局在する分子異常を原因として依存性薬物に対する感受性が上昇することを見出しました (Neuroscience Research, 2020)。このことから、精神疾患病態において、階層回路機構や並列回路機構といった神経回路の制御機構に異常を来とし、報酬/目的指向行動における障害や精神症状につながるとする精神疾患における回路病態仮説を提唱しました (図1, Psychiatry and Clinical Neurosciences, 2019)。今後、モデルマウスの回路病態の解析をすすめて、新しい精神疾患の治療法の開発へとつなげたいと考えています。また、人工知能、ロボティクス、計算精神医学と脳科学の融合に貢献できればと思います。

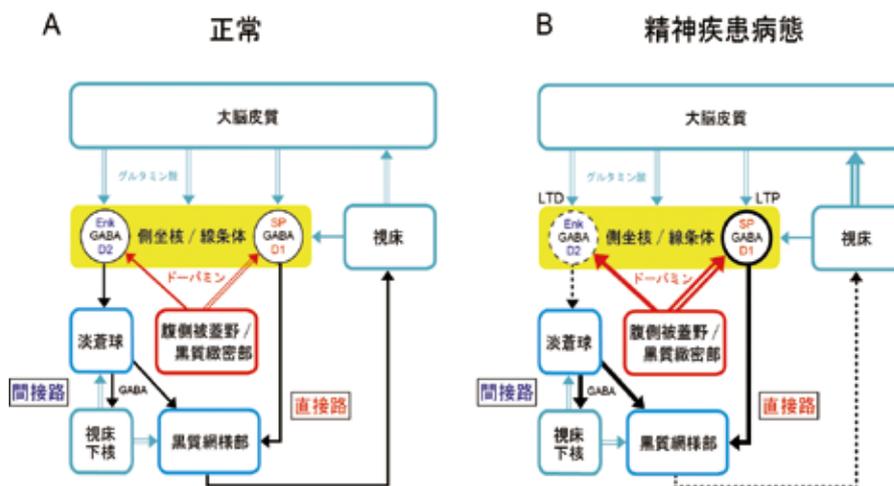


図 大脳基底核神経回路と精神疾患における回路病態仮説

### 主な研究業績:

- 1) Macpherson T, Hikida T (2018). Nucleus accumbens dopamine D1-receptor-expressing neurons control the acquisition of sign-tracking to conditioned cues in mice. *Frontiers in Neuroscience*, 12, 418. <https://doi.org/10.3389/fnins.2018.00418>
- 2) Hikida T, Yao S, Macpherson T, Fukakusa A, Morita M, Kimura H, Hirai K, Ando T, Toyoshiba H, Sawa A (2020). Nucleus accumbens pathways control cell-specific gene expression in the medial prefrontal cortex. *Scientific Reports*, 10, 1838. <https://doi.org/10.1038/s41598-020-58711-2>
- 3) Macpherson T, Matsumoto M, Gomi H, Morimoto J, Uchibe E, Hikida T (2021) Parallel and hierarchical neural mechanisms for adaptive and predictive behavioral control. *Neural Networks*, 144: 507-521. <https://doi.org/10.1016/j.neunet.2021.09.009>

## 報酬と注意の情報処理に関するドーパミン神経回路機構

研究代表者：松本 正幸（筑波大学医学医療系 教授）

ドーパミン神経系は報酬情報をコードする神経系としてよく知られています。特に、その活動は強化学習の中で重要な役割を果たす報酬予測誤差信号をコードし、動物の学習を調節する働きがあると考えられてきました。一方、ドーパミン神経系の異常は報酬とは直接関係のない様々な機能を障害することが知られています。新学術領域「人工知能と脳科学の対照と融合」において、我々の研究チームは、ドーパミン神経系がどのようにしてこのように多様な機能を調節しているのか、そのメカニズムを解析してきました。

我々が注目した脳機能の一つに、「行動抑制」があります。社会生活を送る上で、衝動的な行動や不必要な行動を抑制できることがとても重要です。注意欠陥多動性障害やパーキンソン病などの精神・神経疾患ではこの行動抑制の能力が低下しています。これらの疾患の多くではドーパミン神経系に異常が見られることが知られていますが、この神経系がどのようにして衝動的な行動や不必要な行動を抑制しているのかは全く明らかになっていませんでした。我々のチームでは、行動を抑制することが求められる認知課題をヒトに近縁なマカク属のサルに訓練し、課題遂行中のサルの黒質緻密部および腹側被蓋野のドーパミンニューロンから活動を記録しました (Ogasawara et al., 2018)。実験の結果、サルが行動を抑制することを求められたとき、ドーパミンニューロンの中でも黒質緻密部に分布するものだけが活動を上昇させました。また、黒質緻密部のドーパミンニューロンから投射を受ける線条体領域（尾状核）からも、同様の神経活動の上昇が観察されました。さらには、この線条体領域へのドーパミンニューロンからの神経入力を薬理的に遮断すると、不適切な行動を抑制するサルの能力が著しく低下しました。以上の結果から、黒質緻密部のドーパミンニューロンから線条体尾状核に対して、不適切な行動を抑制するための神経シグナルが伝達されていることが明らかとなりました。

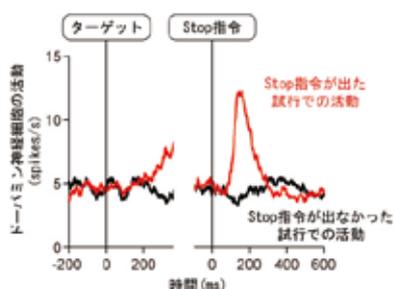


図1 Stop 指令が出た際の黒質緻密部ドーパミンニューロンの活動

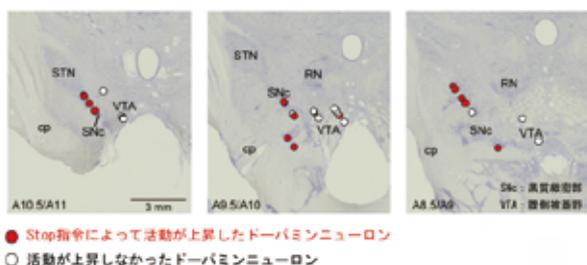


図2 Stop 指令に対して活動が上昇したドーパミンニューロンの分布

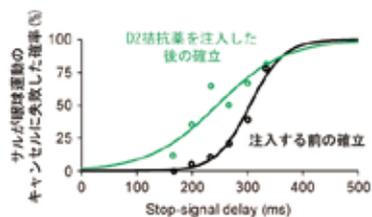


図3 D2 拮抗薬を線条体尾状核に注入した際のサルの行動成績

我々が注目するもう一つの脳機能に「意思決定」があります。我々は日常生活の様々な場面で選択を迫られますが、薬物依存症や強迫性障害では、不合理な意思決定がしばしば見られます。これらの疾患の主な病因の一つは、ドーパミン神経系の異常です。ドーパミン神経系が意思決定に果たす役割を明らかにするため、価値に基づく意思決定が求められる行動課題をマカクザルに行わせ、課題遂行中のサルのドーパミンニューロンから神経活動を記録しました (Yun et al., 2020)。その結果、選択肢が提示されて、サルがそれを選ぶかどうか決めようとしているとき、ドーパミンニューロンは、選択肢の価値が高いほど強く活動を上昇させました。一方、選択肢の価値に関わらず、サルがその選択肢を選ぶときだけ活動を上昇させるドーパミンニューロンや、選択肢の価値とサルの選択行動の両方を反映した活動を示すドーパミンニューロンも多く見つかりました。ここで得られた特に重要な知見は、ドーパミンニューロンは、選択肢が提示された直後は選択肢の価値を反映して活動しますが、選択肢を選ぶかどうか決めようとしている間に、選択行動を反映した活動に変化することです。これまで意思決定の中核として注目されてきた前頭葉の眼窩前頭皮質からも神経活動を記録しましたが、本研究により、ドーパミンニューロンの方が、より早く価値から選択への活動変化を生じることがわかりました。以上のことから、意思決定のメカニズムとして、選択肢の「価値情報」をその選択肢を選ぶかどうかの「選択指令」へといち早く変換しているのは、霊長類で最も発達した高次中枢である前頭葉ではなく、進化的に保存された中脳のドーパミンニューロンであることが明らかになりました。



図4 意思決定時のドーパミン活動ダイナミクス

我々の研究は、ドーパミンニューロンが報酬とは直接関係のない情報をコードし、様々な場面で動物の行動を支えていることを示唆します。現在、計算論的神経科学の研究者らと共同研究を進め、ドーパミン神経系が行動抑制や意思決定を制御する計算モデルの構築を目指しています。最後になりましたが、我々の研究をサポートして頂いた新学術領域「人工知能と脳科学の対照と融合」にお礼を述べて結びとさせていただきます。

### 主な研究業績：

- 1) Ogasawara T, Nejime M, Takada M, Matsumoto M (2018). Primate nigrostriatal dopamine system regulates saccadic response inhibition. *Neuron*, 100, 1513-1526
- 2) Yun M, Kawai T, Nejime M, Yamada H, Matsumoto M (2020). Signal dynamics of midbrain dopamine neurons during economic decision-making in monkeys. *Science Advances*, 6, eaba4962

## 自己と他者の動作データからの内部モデルの構築と行動則の獲得

研究代表者：森本 淳（京都大学 大学院情報学研究科 教授 / 株式会社国際電気通信基礎技術研究所（ATR）室長）

本研究では、脳に学んだ計算原理に基づいて自己や他者の動作データから身体や外界の内部モデルを構築するロボット学習アルゴリズムの開発を目指しました。制御対象のモデルが既知（つまり計算機内に表現可能）であり、時間の制約がなくオフラインでの制御器の導出が許される場合は、モデル予測制御を用いたアプローチが多自由度ロボットの制御に有用であることがわかってきていました。しかし、実環境でのロボットのリアルタイム制御をおこなう場合においては、特に衝突や接触などの実環境を反映したモデルを計算機内に表現することがむずかしく、また制御周期内に制御器導出のための最適化計算が収まらないなどの問題があり、一般には上記アプローチのそのままの活用は容易ではありませんでした。本研究においては、脳の階層的な制御システムを参考に、上位階層においてモデル予測制御により多様な動作目的にあわせた制御器を導出することを可能と

しながら、中間の階層において（モーションキャプチャした）人の動作データからパターン生成器を構成、動作生成の探索範囲を誘導、下位階層においては短い制御周期での計算が可能な目標軌道追従制御を導入することを提案しました。これによって、上位階層での制御周期を長く取れるようになり、ヒト型ロボットモデルにおいてリアルタイムで多様な動作生成が可能となりました。下位階層での柔軟な関節制御の実現は実システムへの実装過程における接触・衝突の扱いの難しさを緩和します。学習試行を通じてその柔軟さを適応させながら、モデル予測制御に用いる内部モデルを更新する当初目標とした方法論の開発にまでたどり着きました。本領域内での神経科学分野の研究者の方々との議論を通じてロボットの全身運動学習手法を発展させることができました。本領域において上記の研究を支援いただいたことに感謝いたします。

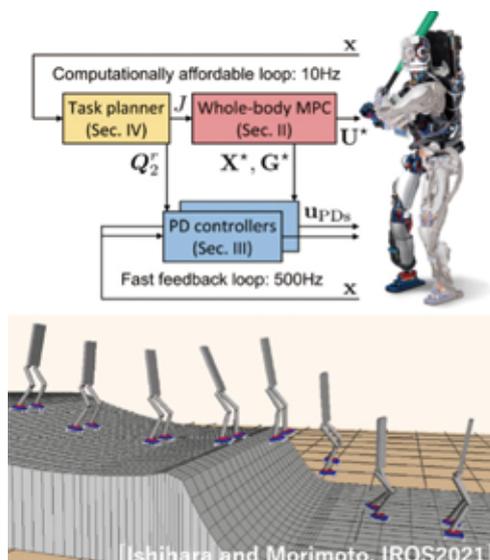


図 脳に学んだ階層的かつ適応的な学習制御

### 主な研究業績：

- 1) Morimoto J (2017). Soft Humanoid Motor Learning. Science Robotics (Focus), vol. 2, issue 13.  
<https://doi.org/10.1126/scirobotics.aaq0989>
- 2) Ishihara K, Itoh T, Morimoto J (2019). Full-Body Optimal Control Toward Versatile and 1Behaviors in a Humanoid Robot. IEEE Robotics and Automation Letters, Vol. 5, Issue 1, 119-126.  
<https://ieeexplore.ieee.org/document/8865637>
- 3) Ishihara K, Morimoto J (2021). Computationally Affordable Hierarchical Framework for Humanoid Robot Control. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2021).  
<https://ieeexplore.ieee.org/document/9636013>

## 感覚予測と報酬予測を統合するメタ学習機構：計算論的理解と脳内基盤の解明

研究代表者：井澤 淳（筑波大学 システム情報系 准教授）

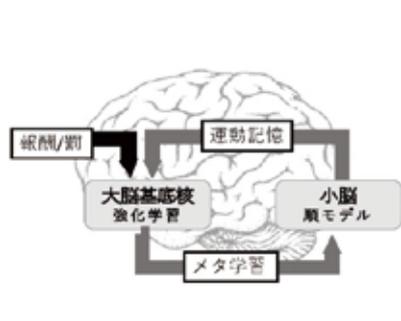
運動学習能力は先天的であり、これにより運動の得手不得手が決まると、一般には信じられています。しかし、実際には運動学習能力は多様に変化します。例えば、トレーニング2日目の学習スピードのほうが1日目の学習スピードよりも速くなる現象が多数報告されています (Krakauer, JNS, 2005)。また、学習スピードが環境変動 (Castro, Cur Biol. 2014, Herzfeld, Science, 2015) や金銭フィードバック (Galea, NatureNeuro, 2015) によっても多様に変化することが、近年明らかになってきました。それぞれ、記憶の想起、誤差の統計的性質、誤差の脳内表現、モチベーションの変化など、様々な理論が提案されましたが、互いに矛盾しているため、背景にある脳内機序も未解明です。また、これら従来理論はいずれも外的要因に影響を受ける受動的な学習能力変容を想定しているため、脳が運動学習能力を能動的に調節する方法は明らかにされていませんでした。

そこで我々は最近大きく注目を集めている人工知能のメタ学習理論（「学習の学習」という階層構造によりメタ学習を実現, Duan 2016, Finn 2017, Wang 2017）に着目し、これをヒトに対する運動学習課題に適用しました。代表的な運動適応課題である視覚運動回転タスクを行っている際に、各予測誤差の経験に対する修正動作に対して、その修正動作の程度に応じて金銭的報酬を与える実験を行いました。1つ目のグループには修正動作が大きいと高い金銭報酬を与え、2つ目のグループには修正動作が小さいと高い金銭報酬を与えました。したがって、被験者が金銭報酬を最大化させるためには、強化学習を通じて能動的に運動適応ゲインを調節する必要があります。実験の結果、1つ目のグループでは運動適応ゲインが大きくなり、2つ目のグループでは運動適応ゲインが小さく学習されました (図B)。この結果から、ヒトの運動適応メカニズムには、強化学習を通じたメタ学習機能が備わっていることが明

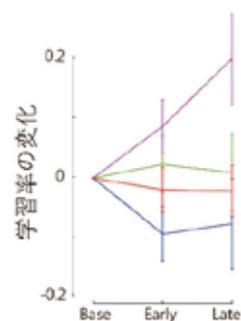
らかになりました<sup>1)</sup>。計算論的には、観測した予測誤差から運動の修正までの感度係数を強化学習アルゴリズムで学習していることを示唆しています。

さらに、我々は、この感度係数が、複雑な身体—環境に対してどのようにメタ学習されるかを調べました。データグローブから任意の変換マトリクスを通じて画面上のカーソルを操作するタスクを実施しました。計算論的に考えれば、このタスクでは、手指の姿勢とカーソルとの関係に冗長な自由度があり、その構造を脳内にモデル化し、そのモデルに基づいた誤差のフィードバックを行う必要があります。そして、一定のトレーニングが終了した段階で、カーソルに外乱を与え、観測した誤差に基づいた運動の修正ゲインを推定しました。その結果、運動学習時の軌道ノイズが修正ゲインを因果的に調節していることが明らかになりました。これは修正ゲインの設計に用いる、手指の姿勢とカーソルとの関係の構造の学習に関して軌道ノイズが効果的である事を意味しており、予測誤差に基づいた運動学習における探索ノイズの役割を示唆しています<sup>2)</sup>。これまで報酬に基づいた運動学習における探索ノイズの役割は発見されていましたが、予測誤差に基づいた運動学習における探索ノイズの役割は初めての報告でした。

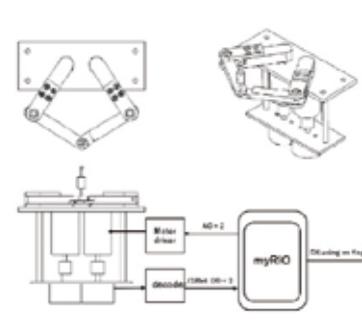
これら研究のように、運動制御・学習の計算論は、計算機実験の結果とヒト行動実験との比較によって行われてきました。一部、運動制御中の脳活動を脳機能イメージングで計測した研究もありますが、スキャナー中には実験装置（ロボットマニピュラタム等）の使用が制限されており、行動実験と比べると単純な運動学習実験に限定されていました。そこで、メタ運動学習中の全脳の活動を調べ、その脳内実装の全貌を明らかにするための第一歩として、げっ歯類用のロボットマニピュラタムを開発しました (図C)<sup>3)</sup>。これは当該新学術領域内の共同研究としてスタートしたプロジェクトです。



図A 運動メタ学習アーキテクチャ



図B メタ学習効果 (ヒトに対する実験の結果)



図C げっ歯類用マニピュラタム

### 主な研究業績：

- 1) Sugiyama, T., Schweighofer, N., & Izawa, J. (2020). Reinforcement meta-learning optimizes visuomotor learning. *bioRxiv*. <https://doi.org/10.1101/2020.01.19.912048>
- 2) Dal' Bello, Lucas Rebelo, and Jun Izawa. Dal' Bello, L. R., & Izawa, J. (2021). Task-relevant and task-irrelevant variability causally shape error-based motor learning. *Neural Networks*, 142, 583-596. *Neural Networks* 142 (2021): 583-596. <https://doi.org/10.1016/j.neunet.2021.07.015>
- 3) Inoue, T., Terada, S. I., Matsuzaki, M., & Izawa, J. (2021). A small-scale robotic manipulandum for motor control study with rodents. *Advanced Robotics*, 1-9. <https://doi.org/10.1080/01691864.2021.1912637>

## 霊長類の前頭前野を巡る「行動選択」神経ネットワークの構築様式 — 神経解剖学および自由行動解析への機械学習手法の導入 —

研究代表者：井上 謙一（京都大学霊長類研究所 助教）

私達は、脳型 AI のデザインに資する大脳皮質と皮質下構造との神経連絡様式を解明することを目的に、多重越シナプストレーシング法や、さらに順行性軸索トレーシングと組み合わせた双方向性トレーシング法（図 A）を開発し、大脳皮質と大脳基底核・扁桃体・小脳などとの連絡回路を解析してきました<sup>1)2)</sup>。また、私自身はこれまで AI に関しては全くの素人でしたが、領域活動や共同研究などを通じ、AI 技術を自身の脳研究へ利用する試みを進めてきました。本稿ではその中の 2 例について紹介します。

### 霊長類神経トレーシングデータの自動ラベリング

霊長類における多重トレーシング実験では、大きな脳で多重標識細胞を一つずつ検出していくのには膨大な時間がかかるという問題を抱えていました。これを解決するため、私達は機械学習を用いて標識された神経細胞を自動で識別するパイプラインの開発に取り組みました。マカザルの脳はマウスの約 200 倍もの大きさを持つため、データサイズの限界から 1 細胞あたりの像は 10x10 ピクセル程度となってしまう、また越シナプスラベルは領域によるラベルの強さや密度が大きく異なりますが、そのような条件でも精度良いラベル検出を実現することを目指しました。その結果、研究者が行うラベル検出と精度が同等であるネットワークを構築することができました。また、転移学習により、別の個体や別種のウイルスベクターを注入したサンプルに対しても、新たな個体の 500-1000 細胞分のみ教師データを加えることによって、十分な精度での検出が可能であることが分かりました

（図 B）。現在、多重蛍光ラベル細胞の検出のためのシステム改良を行っており、今後全脳におけるラベルをデジタルデータ化することにより、複数の実験のデータ統合による包括的解析や、データベース化などに寄与できるのではないかと期待しています。

### マカサルモーションキャプチャーシステムの開発

構造的にも機能的にもヒトと類似した脳を持つサル類は、疾患の治療法開発やヒトの高次脳機能研究に必須の動物モデルです。言語を有しないサルにおいては、姿勢や動作の解析が、運動機能だけでなく社会性および情動性などの評価に必要となり、サルの姿勢や動作を定量的かつ高い再現性で解析可能なマーカーレスモーションキャプチャー技術の重要性は高いと考えられます。そこで私達は、富士大学の松本助教と九州工業大学の柴田教授と共同で、まず AI にサルの体の特徴点（各関節や目鼻耳など）の位置を学習させるため、約 17,000 頭分のマカサルの画像を収集してラベル付けを行うことで、大規模かつ高品質なサル姿勢の教師データセットを作成しました。そして、転移学習による効率的な学習を行うことができるアルゴリズムを適用して、その推定精度を検証しました<sup>3)</sup>。その結果、約 90% の精度で特徴点が検出でき、特徴点の位置の誤差も人が行った場合と同等と、高い精度で特徴点推定が可能であることが分かりました（図 C）。現在これを発展させて、3 次元姿勢の推定や複数個体の姿勢推定を実現するアルゴリズム開発を行うとともに、非ヒト霊長類の自然行動解析への適用研究を進めています。

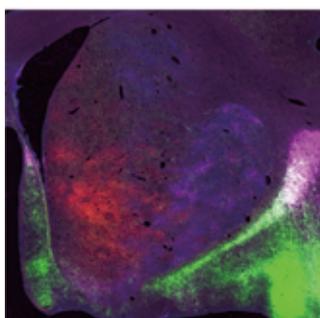


図 A 双方向性トレーシング法の実施例。青・赤色が軸索ラベル、緑・桃色が逆行性ラベル。

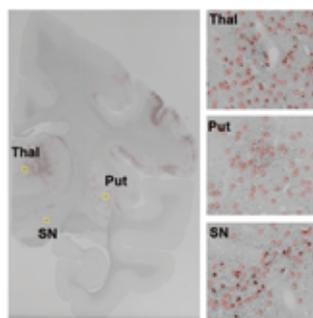


図 B AIを利用した逆行性標識ニューロンの自動ラベル。赤丸がラベルされた標識ニューロン。Tal: 視床、Put: 線条体、SN: 黒質。



図 C AIを利用したマカサルモーションキャプチャ<sup>3)</sup>。

#### 主な研究業績：

- 1) Ishida H<sup>#</sup>, Inoue K<sup>#</sup>, Takada M. <sup>#These authors contributed equally.</sup> (2018) Multisynaptic projections from the amygdala to the ventral premotor cortex in macaque monkeys: Anatomical substrate for feeding behavior. *Front Neuroanat*. 12: 3. doi: 10.3389/fnana.2018.00003.
- 2) Lu X, Inoue K, Ohmae S, Uchida Y. (2020) New Cerebello-Cortical Pathway Involved in Higher-Order Oculomotor Control. *Cerebellum* 19:401-408. doi: 10.1007/s12311-020-01108-8.
- 3) Labuguen R, Matsumoto J, Negrete S, Nishimaru H, Nishijo H, Takada M, Go Y, Inoue K, Shibata T. (2021) MacaquePose: A novel 'in the wild' macaque monkey pose dataset for markerless motion capture. *Front Behav Neurosci*, 14:581154. doi: 10.3389/fnbeh.2020.581154.

## モデルフリーモデルベースの協同による深層並列強化学習

研究代表者：内部 英治（国際電気通信基礎技術研究所 主幹研究員）

行動学習の一つである強化学習は環境から得られる「状態・行動・報酬」の系列から直接方策を学習するモデルフリー法と、いったん状態遷移確率を生成モデルとして学習するモデルベース法があります。両者は状況に応じて使い分けられていることが示唆されていますが、その基準は不明でした。本研究では、価値関数を評価基準としてモデルフリー・モデルベースを選択する並列学習法を、意思決定に要する時間の差を考慮するように拡張した非同期方法を開発しました（図 a）。MuJoCo 課題で検証したところ、学習初期の探査段階ではモデルフリー、学習中期ではモデルベース、学習後期では再びモデルフリーが自

動的に選ばれることを確認しました（図 c）。また、時間の差を考慮しない同期した協同方法よりも、非同期の協同方法が学習時のデータ効率も改善できることも確認しました（図 b）。

また正の報酬の最大化と罰の最大化を個別に並列学習する MaxPain 法の学習過程を安定化させるために、報酬最大化方策と罰最大化方策の間の Kullback-Leibler ダイバージェンスを正則加項として追加する方法を開発しました（図 d）。移動ロボットのナビゲーション課題（図 e, f）による検証では、従来よりも統合に関するハイパーパラメータの調整が容易で、制御性能もロバストになる結果が得られました（図 g）。

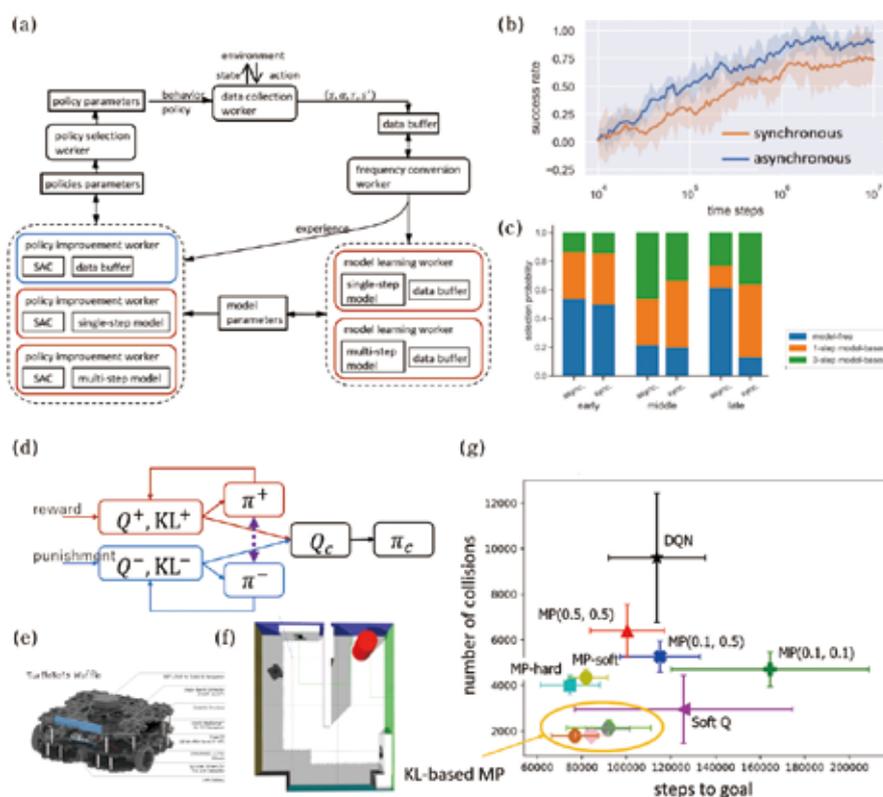


図 モデルフリーモデルベースの協同による深層並列強化学習

a: 非同期かつ協調的に動作する深層並列強化学習のアーキテクチャ。 b: 非同期に協調すると同期して強調する場合よりも学習効率が改善される。 c: 学習過程でのモデルフリー・モデルベース強化学習の選択の推移。 d: 報酬と罰を分離して学習する MaxPain アーキテクチャ。 e: 実験に用いた移動ロボット。 f: U-maze 環境。 g: MaxPain はハイパーパラメータの選択に影響を受けにくい。

### 主な研究業績：

- 1) Wang, J, Elfving, S, Uchibe, E (2021). Modular Deep Reinforcement Learning from Reward and Punishment for Robot Navigation. *Neural Networks*, 135, 115-126.
- 2) Uchibe, E, Doya, K (2021). Forward and Inverse Reinforcement Learning Sharing Network Weights and Hyperparameters. *Neural Networks*, 144, 138-153.
- 3) Macpherson, T, Matsumoto, M, Gomi, H, Morimoto, J, Uchibe, E, Hikida, T (2021): Parallel and Hierarchical Neural Mechanisms for Adaptive and Predictive Behavioral Control, *Neural Networks*, 144, 507-521.

## ニューラルネットワークによる神経ネットワークの動作原理の解明

研究代表者：飯野 雄一（東京大学 大学院理学系研究科 教授）

私達は神経回路の構造がすべてわかっている小動物線虫を用いて行動を生み出す神経系の動作原理を解明することを目指しました。まず、多数の線虫の行動をcMOSカメラで測定して線虫の刻々の姿勢を数値化し、得られたデータを再帰的ニューラルネットワークRNNに混合ガウス分布による確率モデルを組み合わせたMDN-RNNニューラルネットワークモデルに学習させることにより、線虫の行動をよく再現するモデルが作成でき、それぞれのガウス分布が特定の行動シーケンスに対応していることが分かりました。このニューラルネットワークの学習を用いることにより自動的に行動分類が達成できることが明らかとなりました。

一方、対物レンズを高速に移動させて頭部全神経を観測する4Dイメージング顕微鏡により神経活動を測定し、得られた約200神経の活動時系列データをモデル化しました。ここでも非線形次元削減に引き続き混合分布の確率モデルを用いることにより、実際の神経活動のパターンを、確率的でありつつ神経間の相互関係をよく保ったままでモデル化することができました。モデル上での神経破壊などによりネットワーク上の情報の流れが推定できました。

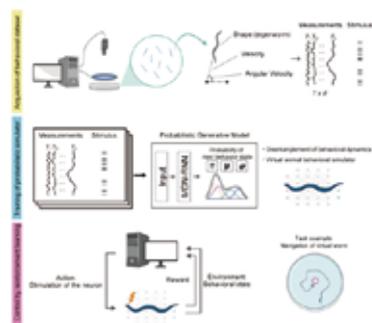


図1 MDN-RNNによる線虫の行動のモデル化  
神経の行動をカメラで経時的に観測して数値化し、MDN-RNNで時系列を学習させることにより確率的行動をよくモデル化できた。さらには得られたモデルを用い強化学習で行動制御の手法を獲得できた。

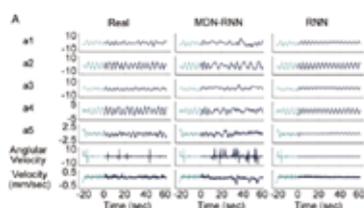


図2 多彩な確率的行動の再現にはMDN-RNNモデルが適し、RNNモデルでは不十分。

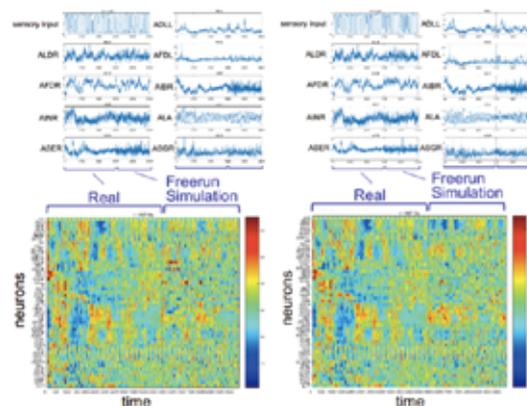


図3 神経活動のモデル化  
混合ガウス分布による確率モデルでは繰り返しシミュレーションを行うと異なる結果となるが神経間の関係はよく再現されている。

### 主な研究業績：

- 1) Mori K, Yamauchi N, Wang H, Sato K, Toyoshima Y, Iino Y (2021). Probabilistic generative modeling and reinforcement learning extract the intrinsic features of animal behavior. *Neural Networks*, 145, 107-120. <https://doi.org/10.1016/j.neunet.2021.10.002>
- 2) Sato H, Kunitomo H, Fei X, Hashimoto K, and Iino Y (2021). Glutamate signaling from a single sensory neuron mediates experience-dependent bidirectional behavior in *Caenorhabditis elegans*. *Cell Rep* 35, 109177. 10.1016/j.celrep.2021.109177.
- 3) Toyoshima Y, Wu S, Kanamori M, Sato H, Jang MS, Oe S, Murakami Y, Teramoto T, Park C, Iwasaki Y, Ishihara T, Yoshida R, Iino Y (2020). Neuron ID dataset facilitates neuronal annotation for whole-brain activity imaging of *C. elegans*. *BMC Biol* 18, 30.

## 刺激弁別の学習フェーズにおける線条体細胞タイプ特異的な神経活動の可視化

研究代表者：小林 和人（福島県立医科大学 生体機能研究部門 教授）

複雑な運動技能や意思決定の基盤となる学習プロセスは、環境内の刺激と適切な行動の間の連合を形成することであり、皮質と大脳基底核を結ぶ神経ネットワークによって支えられていると考えられています。大脳基底核の入力層に位置する線条体は、出力層への投射様式の違いに基づいて分類された直接路ニューロンと間接路ニューロンから構成されます。この2タイプの神経細胞が、刺激-行動学習プロセスにおいてユニークな役割をはたすことは、我々の研究グループの成果を含む過去の研究結果から示唆されてきましたが、その全容解明には至っていません。

そこで我々は、細胞機能を操作するための遺伝子を、直接路ニューロンあるいは間接路ニューロンに特異的に発現誘導することを可能にする遺伝子改変ラット（Tac1-Cre/Drd2-Creラッ

ト）を作製しました。これらのラットとウイルスベクター技術に基づき、直接路/間接路ニューロンの選択的な除去（緑膿菌外毒素・PE38）、活動可視化（カルシウム感受性蛍光たんぱく質・GCaMP）、活動操作（昆虫由来の陽イオンチャネル・IRNA）のためのシステムを確立しました。今後はこれらのシステムを有機的に組み合わせ、刺激-行動学習における様々な学習フェーズにおける2タイプの神経細胞のはたらきとその意義について、AIの学習と対照しつつ解明したいと考えています。

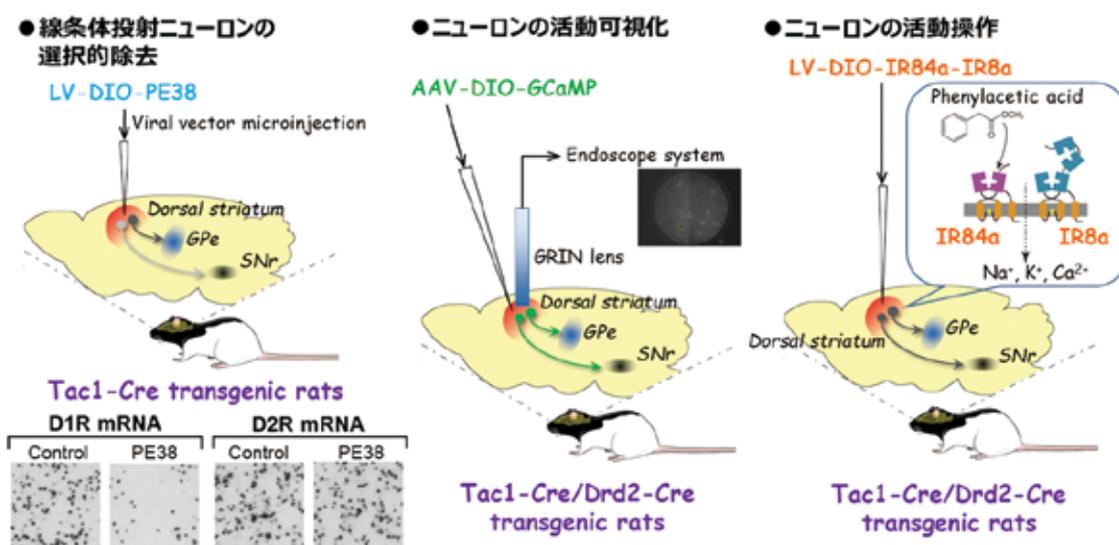


図 ラットの学習過程における線条体細胞タイプ（直接路/間接路ニューロン）特異的な役割を解明するために独自に開発した分子遺伝学的ツール。左から、緑膿菌外毒素を用いた選択的除去・カルシウム感受性蛍光タンパク質を用いた活動可視化・昆虫由来の陽イオンチャネルを用いた活動操作（亢進）。

### 主な研究業績：

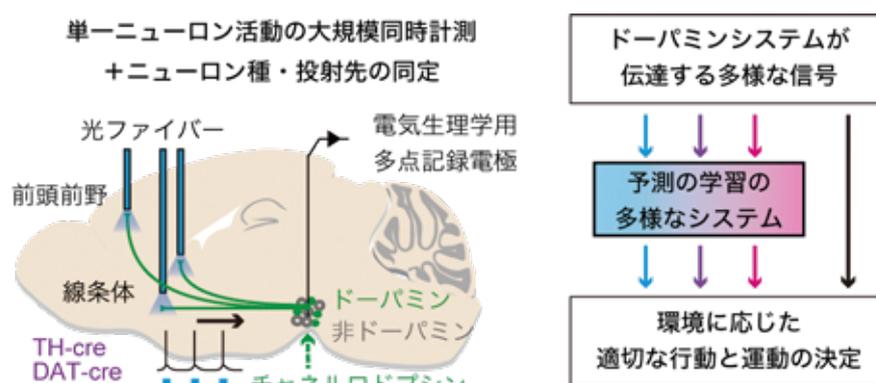
- 1) Osanai M, et al. (2021). Multimodal functional analysis platform: 1. Ultrathin fluorescence endoscope imaging system enables flexible functional brain imaging. In: Yao H. (ed) Optogenetics. Advances in Experimental Medicine and Biology, 1293: 471-479. Springer Nature Singapore Pte Ltd.
- 2) Fukabori R, Iguchi Y, et al. (2020). Enhanced retrieval of taste associative memory by chemogenetic activation of locus coeruleus norepinephrine neurons. Journal of Neuroscience, 40, 8367-8385.

## 行動選択におけるドーパミン神経回路の並列情報処理機構の解明

研究代表者：松本 英之（大阪市立大学大学院医学研究科 神経生理学 助教）

未来を予測して行動し、その行動の結果と予測の誤差からさらに良い予測を学ぶ能力は、環境に応じた適切な行動と運動の選択にとっても役に立ちます。中脳ドーパミン細胞は、報酬に基づく予測誤差学習に重要な役割を果たすことが知られています。一方で近年の研究から、ドーパミン細胞にはさまざまな多様性があることが明らかになってきました。このことから、ドーパミン細胞はこれまで考えられていた以上に多様で広範な学習に関与している可能性が予想されます。本研究では、ドーパミン細胞にみられる多様性のうち、投射神経回路の違いと環境依存的な信号モードの変化に着目し、価値に基づく行動選択、意思決定との関連を調べました。大規模な単一細胞活動記録法、イ

メージング法、光遺伝学を組み合わせた研究技術を、試行錯誤を通して運動や行動を柔軟に変化させる自由行動中の動物に適用し、異なる脳領域へと伝達されるドーパミン信号をさまざまな時間・空間のスケールで計測しました。その結果、投射先別のドーパミン予測誤差信号は、その性質に違いがあることがわかりました。また、環境に適応する過程において、ドーパミン信号は多様な時間スケールで修飾されていることがわかりました。ドーパミン神経系は、投射先別の異なる学習信号を伝達し、信号モードを柔軟に修飾させるメカニズムを通して、環境に適した行動や運動の出力に効果的に関与している可能性が考えられます。



図左 投射先別のドーパミン信号を高時間分解能・単一細胞レベルで計測する技術の概要。中脳にあるドーパミン神経起始核から電気生理学的手法を用いて単一ニューロン活動を大規模に同時計測する。ドーパミン神経細胞を光感受性陽イオンチャンネルであるチャンネルロドプシンで特異的に標識する。ドーパミン神経細胞の投射先と起始核に刺入された極細の光ファイバーを通して、ドーパミン神経細胞の軸索あるいは細胞体に直接光を照射することで、チャンネルロドプシン陽性のドーパミン神経細胞に活動電位を人為的に引き起こす。この活動電位の特徴を解析することで、大規模に同時計測されたニューロンのうちのどれがドーパミン神経細胞種で、どの脳領域に投射しているのか、同定することができる。この技術を認知行動課題中の動物に適用して、課題中の投射先別ドーパミン信号を高時間分解能で計測した。

図右 ドーパミン神経細胞は投射先別の異なる学習信号を伝達している可能性が考えられる。ドーパミン信号のこのような多様性は、適切な行動や運動の決定において、別の脳領域にある異なる学習システムの個別最適化に都合が良いと考えられる。

### 主な研究業績：

- 1) Matsumoto H, Mizuseki K (2021). Projection-specific roles of midbrain dopamine neurons in reward-based decisions. The 44th Annual Meeting of the Japan Neuroscience (Neuroscience 2021)
- 2) Tsutsui-Kimura I, Matsumoto H, Akiti K, Yamada MM, Uchida N, Watabe-Uchida M (2020). Distinct temporal difference error signals in dopamine axons in three regions of the striatum in a decision-making task. *Elife* 9, e62390  
<https://elifesciences.org/articles/62390>
- 3) 松本英之、内田光子、内田直滋 (2020). 「強化学習とドーパミンの多様性」、*ブレインサイエンスレビュー* 2020 (論文集)、77-101、クバプロ

## 前頭前野における情報の抽象化と演繹的創生の神経メカニズムの研究

研究代表者：坂上 雅道 (玉川大学脳科学研究所 教授)

### 課題遂行下におけるマカザル第1次視覚野からの微小内視鏡カルシウムイメージング

これまで、マカザルを含む霊長類からの神経活動記録においては、電極を使った電気生理学的な手法が主に用いられてきました。これは、神経細胞が活動する時の電気的な信号変化を捉えるものです。一方で、神経細胞が活動する時には、細胞内部でのイオン濃度の急激な上昇も生じます。このうち、カルシウムイオン濃度の変化を蛍光信号に変換し、多数の神経細胞の活動を同時に可視化する技術が「カルシウムイメージング」です。本研究では、イメージング法として、脳内に微小な筒形のレンズを埋め込み、ミニチュア顕微鏡を使って観察する「微小内視鏡法」を用いました。微小内視鏡法は、脳深部からの神経活動記録や自由行動下における記録に適しているというメリットがあります。微小内視鏡法はこれまで主にマウスなどで用いられてきましたが、技術的なハードルから、霊長類への適用例はほとんどありませんでした。

私たちは、マカザルの第1次視覚野 (V1) をターゲットとして、微小内視鏡法を適用するための方法を開発しました。カルシウムイメージングのためのセンサーを神経細胞に合成させるため、公募班の井上謙一助教 (京都大学) の協力の元、新たに開発された、霊長類で高い感染効率を誇るアデノ随伴性ウイルスベクター 2.1 型 (AAV2.1) を用いました。3頭のサルに対して、両側 V1 へのウイルスの注入後、GRINレンズの留置と固定を行いました。サルは、イメージング記録中、視野周辺に特定の傾きをもった視覚刺激が一瞬呈示される、簡単な注視課題を行いました。

実験の結果、2頭のサルの3つの大脳半球から、視覚刺激に対して応答を示す多数の神経細胞を観察することができました (図1)。これらの映像から神経細胞を検出し、そのデータを用いて解析を行ったところ、観察された神経細胞から V1 細胞がもつ受容野特性や傾き選択性といった特徴が確認できました。また、サルが見ている視覚刺激の傾きを機械学習の方法を使って解読する (デコーディングすること) や、異なる日に得られたデータを使った細胞追跡にも成功しました。以上の結果は、微小内視鏡カルシウムイメージング法がマカザルなどの霊長類においても有効な観察法として利用可能であることを実証するものであり、微小内視鏡法を用いた他の脳領域における局所的な神経回路のダイナミクスの解析へも通ずるものです。<sup>1)</sup>

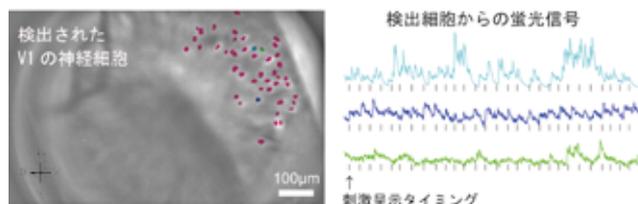


図1 マカ V1 からの微小内視鏡カルシウムイメージング

### 化学遺伝学2重遺伝子導入法を用いたマカ前頭前野 - 線条体経路の機能解明

私たちの脳が果たすさまざまな機能は、膨大な数の神経細胞によって織りなされる複雑な神経ネットワークを通じて実現されています。神経回路と脳機能との複雑な関係を解きほぐすためには、特定の神経回路に対してのみ影響を及ぼす選択的な手法が必要となります。ところが、これまでの電気刺激や薬物注入を用いた操作は、介入部位を通るさまざまな神経回路に対して無差別的に作用するものでした。近年、遺伝子工学的な技術の進展により、経路選択性をもった介入手法がげっ歯類での研究を中心に盛んになってきています。しかし、よりヒトとの近縁性の高い霊長類ではまだ十分に適用が進んでいません。

本研究では、特殊な薬剤 (CNO) を投与することによって特定の経路を担う神経細胞のみを操作することを可能とする「化学遺伝学2重遺伝子導入法」という手法を用い、マカザルの外側前頭前野 (LPFC) から線条体尾状核 (CdN) へ投射する経路を制御することでその機能を調べました (図2左)。このLPFC-CdN経路は、意思決定において重要な役割を果たすと考えられており、ヒトのイメージング研究などによって「抑制コントロール機能」や「ワーキングメモリー機能」との関連が示唆されてきました。しかし、この経路を対象とした、経路選択的な操作を用いた因果的なレベルでの解明はこれまで手つかずの状態でした。

実験では、2頭のサルのLPFC-CdN経路に対してこの化学遺伝学2重遺伝子導入法を用いました。これらのサルには、あらかじめ、抑制コントロールとワーキングメモリーへの影響を同時に検証可能な課題として、報酬が非対称的に配分されるサッケード遅延反応課題を訓練しました。課題遂行下のサルにCNOを投与したところ、コントロール条件と比べ、投与後に連続でエラーをおかす頻度が増え、最後まで課題を完遂できないセッションの比率が大きくなりました。また、サルが報酬を得るために行う眼球運動では、潜時が短くなるなど衝動性の亢進を示す結果が得られました。また、記録した局所場電位の解析からは、CNO投与下で課題に関連する神経活動が減弱化することが確認できました (図2右)。これらの結果はLPFC-CdN経路の制御が抑制コントロール機能の低下をもたらすことを示唆しています。この成果は、霊長類を用いたより精密な神経回路研究への歩みを前進させるものであり、当該経路との関係が指摘されているさまざまな精神疾患の病態解明や新規治療法の開発へも結びつくことが期待されます。<sup>2)</sup>

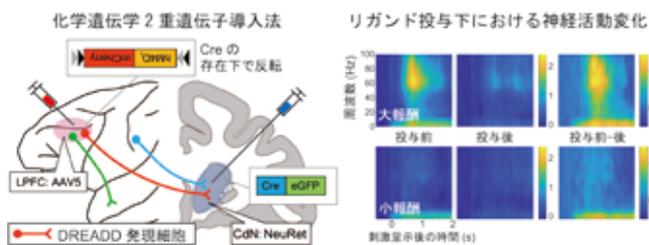


図2 マカ前頭前野 - 線条体経路の化学遺伝学による選択的操作

#### 主な研究業績：

- 1) Oguchi, M. et al. (2021). Microendoscopic calcium imaging of the primary visual cortex of behaving macaques. *Scientific reports*, 11(1), 17021. <https://doi.org/10.1038/s41598-021-96532-z>
- 2) Oguchi, M. et al. (2021). Chemogenetic inactivation reveals the inhibitory control function of the prefronto-striatal pathway in the macaque brain. *Communications biology*, 4(1), 1088. <https://doi.org/10.1038/s42003-021-02623-y>

## 精神疾患における思考の障害の神経基盤の解明と支援法の開発

研究代表者：高橋 英彦（東京医科歯科大学大学院医歯学総合研究科 精神行動医学 教授）

統合失調症は、妄想や支離滅裂な言動などの思考障害を特徴とする精神疾患です。20世紀初頭にE. Bleulerによってその疾患概念が確立されて以来、思考障害は概念間の意味連関の異常、すなわち「連合弛緩」に起因すると考えられてきました。「連合弛緩」は現代においても重要な精神病理ですが、その病態生理はこれまで未解明でした。

本研究では、統合失調症患者の脳活動において概念間の意味連関がどのように表象されるかを評価するために、fMRIを用いて意味表象の大規模ネットワーク（脳内意味ネットワーク）の構造的特徴を調べました。まず、エンコーディングモデルを用いて、動画を見たときの脳活動から患者の脳内における様々な意味表象を定量化しました。続いて、これらの意味表象の類似性によって脳内意味ネットワークを構築し、グラフ理論に基づくネットワーク解析を行いました。

その結果、スモールワールド性は、統合失調症患者の脳内意味ネットワークにおいて、健常者の脳内意味ネットワークや自然言語ネットワークよりも有意に低下しており、妄想の重症度と負の相関が認められました。また、患者の意味ネットワークは、健常者よりも明確にカテゴリーに区分されていましたが、各カテゴリーの内部構造はランダム化していました。これらの結果は、統合失調症の脳内意味ネットワークにおける局所クラスタリング異常と意味連関の侵入性変化を示唆します。本研究は、統合失調症における脳内意味ネットワークのランダム化、すなわち脳における「連合弛緩」の様相を、初めて直接的に明らかにしたものです。これらの知見は、精神疾患患者の内的体験へのアプローチを可能にするだけでなく、認知神経科学、神経画像、自然言語処理など多岐にわたる領域において、脳内の知識体系の理解に貢献することが期待されます。

統合失調症の根本的な病態を解明するには、脳における情報処理の素過程と考えられる、局所回路での計算に注目し、その異常を明らかにする必要がありますと考えています。局所回路での計算とは、たとえば、視覚で言えば、V1でのエッジ検出などのことです。しかし、局所回路での処理をとらえるには、個々のニューロン（あるいは少数のニューロンの集合）の活動をとらえる必要があると考えられますが、人間においては、直接電極でニューロンの活動を調べることは通常はできません。fMRIは、侵襲が少なく貴重なものと考えられますが、局所回路の計算過程を知るには空間解像度が障壁となります。そこで、最近開発された方式を用い、fMRIデータからmultivariateに情報を得ることによって、fMRI信号と、情報処理過程を模したニューラルネットワークモデル（私たちが用いたのは、視覚における、AlexNet）の活動を対応づけ、それを利用して、局所回路の計算異常を検出する実験を計画しました。統合失調症の患者群とコントロール群に、まずは、750枚の自然画像を見せ、その時の脳活動をfMRIで記録し、これらのトレーニングデータを用いて、fMRI信号をAlexNetの活動に翻訳できる、デコーダーを作りました。次に、2群の被験者に、ぼやけた画像を見せ、先ほどのデコーダーを使って、その時のfMRI信号から情報処理過程で不完全な情報がどのように補完処理されているのかを知ることを試みたのです。その結果、統合失調症群では、ぼやけた画像の情報処理はコントロール群とは異なったものになっているようであるが、そこから画像の補完がコントロール群よりも大きくなされ、情報が同程度に復元されているようであることがわかりました。現在、統合失調症の局所回路における計算異常に示唆をもたらすべく、統合失調症群でぼやけた画像の処理がどのようになされているのかを解析中です。

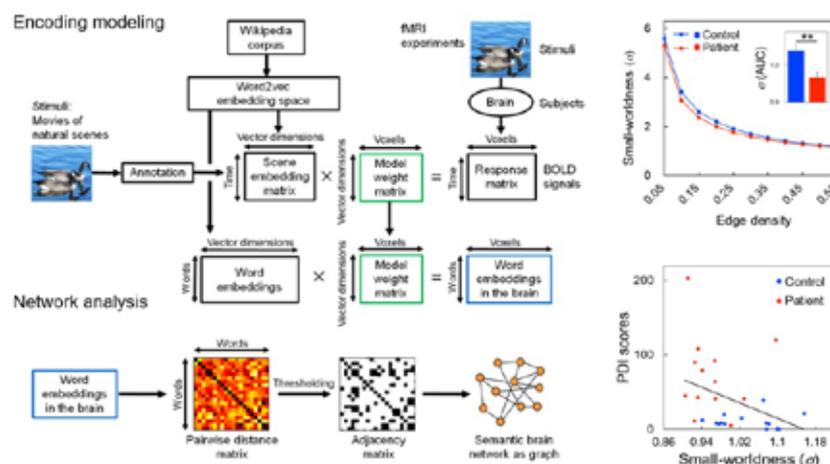


図1 実験手順と、脳内意味ネットワーク指標及び妄想指標との相関

## 主な研究業績：

- 1) Matsumoto Y, Nishida S, Hayashi R, Son S, Murakami A, Yoshikawa N, Ito H, Oishi N, Masuda N, Murai T, Friston K, Nishimoto S, Takahashi H (2021) Disorganization of semantic brain networks in schizophrenia. submitted.
- 2) Nishida S, Matsumoto Y, Yoshikawa N, Son S, Murakami A, Hayashi R, Nishimoto S, Takahashi H (2020) Reduced intra- and inter-individual diversity of semantic representations in the brains of schizophrenia patients. <https://www.biorxiv.org/content/10.1101/2020.06.03.132928v1>
- 3) Yoshihara Y, Lisi G, Yahata N, Fujino J, Matsumoto Y, Miyata J, Sugihara GI, Urayama SI, Kubota M, Yamashita M, Hashimoto R, Ichikawa N, Cahn W, van Haren NEM, Mori S, Okamoto Y, Kasai K, Kato N, Imamizu H, Kahn RS, Sawa A, Kawato M, Murai T, Morimoto J, Takahashi H. Overlapping but Asymmetrical Relationships Between Schizophrenia and Autism Revealed by Brain Connectivity. Schizophr Bull. 2020 Apr 17;46(5):1210-8.

## 感覚運動と言語をつなぐ二重分節解析の脳内計算過程の理解と応用

研究代表者：谷口 忠大 (立命館大学 情報理工学部 教授)

人間の hochi 認知機能を代表するのが言語です。その一方で感覚運動情報は多くのが無意識的に処理され、実行され、低次認知機能を担います。人工知能の歴史を紐解くと高次認知機能に関わる言語を記号的に書き下しトップダウンに与えることで知能を作ろうとした時代がありました。一方で近年の深層学習のような脳に学んだアプローチは感覚運動情報を直接ニューラルネットワークに処理させることで、知的な機能を実現しています。この言語的な表現と、感覚運動情報を繋ぐことは人工知能の大きな問題となります。本研究ではその「界面」に注目しました。つまり言語において感覚運動情報を表象する最小単位である単語（語彙）を形成する脳機能に注目しました。それが二重分節解析の脳内計算過程です。

二重分節構造は人間の音声言語には音声信号が音素に分節化され、また、これが単語へと分節化されるという二段階の分節構造を指し、人間の言語に特有の構造と言われます。この事実は人間の脳が二重分節構造を解析することに秀でた計算能力を持っていることを示唆しています。この脳機能に注目し、その脳内表現を探求するとともに、マルチモーダルな感覚運動情報統合の中で言葉が意味を持つプロセスを構成する計算論モデルを研究しました。

研究は (A) 二重分節解析と動的カテゴリ形成の脳内計算過程を表現する内部モデル計算論の構築、(B) 二重分節解析と動的カテゴリ形成の脳内計算過程の解明、(C) 自律的な言語獲得・運動学習を実現

するロボットの創造の三つの副項目により構成されました。

主たる研究成果の一つは文献<sup>1)</sup>に示した場所概念学習と二重分節構造に着目した語彙獲得を同時に進める確率的生成モデルであり、この手法 SpCoSLAM++ を用いることで、ロボットは事前に辞書を持つことなく、環境と人間との相互作用を通して、語彙を発見し、それを活動空間の位置情報と関連付けることが出来ます (図1)。後にさらに議論を海馬系との関係について発展させることもできました。

また運動学習に関しては文献<sup>2)</sup>において模倣学習と強化学習を確率的生成モデルの観点から融合する研究を行いました。これらの研究を通して、明らかになってきたのは複数の確率的生成モデルを結合し、同時学習させることが、様々なマルチモーダル感覚運動統合や言語と感覚運動情報の関連付けのために重要だということです。これはまさに人間の脳において様々な領域が相互結合しながら学習を進めることに対応します。

そこで文献<sup>3)</sup>では領域内連携 (A01 松尾グループ) により深層生成モデルを含んだ確率的生成モデルを統合し、統合的な認知システムを作るための Neuro-SERKET というフレームワークを開発しました (図2)。この研究はさらに全脳確率的生成モデル (Whole Brain-Probabilistic Generative Model: WB-PGM) というアイデアに発展しました。領域終了後もその研究成果は発展的に展開していく予定です。

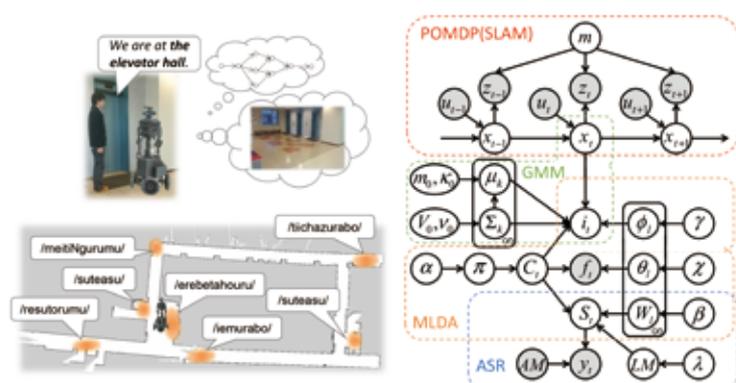


図1 SpCoSLAM++ のグラフィカルモデルと語彙獲得の様子

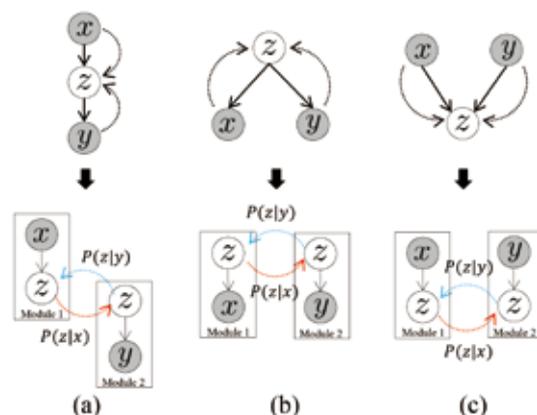


図2 確率的生成モデルを分解しつつ、推論を分散的に行う Neuro-SERKET

### 主な研究業績：

- 1) Taniguchi A, Taniguchi T, Inamura T (2017). Unsupervised Spatial Lexical Acquisition by Updating a Language Model with Place Clues, Robotics and Autonomous Systems, 99, 166-180. <https://doi.org/10.1016/j.robot.2017.10.013>
- 2) Kinose A, Taniguchi T (2020). Integration of imitation learning using GAIL and reinforcement learning using task-achievement rewards via probabilistic graphical model, Advanced Robotics, 34(16),1055-1067. <https://doi.org/10.1080/01691864.2020.1778521>
- 3) Taniguchi T, Nakamura T, Suzuki M, Kuniyasu R, Hayashi K, Taniguchi A, Horii T, Nagai T (2020). Neuro-SERKET: Development of Integrative Cognitive System through the Composition of Deep Probabilistic Generative Models, New Generation Computing, 84. <https://doi.org/10.1007/s00354-019-00084-w>

## 脳内他者を生かす意思決定の脳計算プリミティブの解明

研究代表者：中原 裕之（理化学研究所 脳神経科学研究センター 学習理論・社会脳研究チーム）

私たちの意思決定は、主には自分自身への利益（報酬）に基づいて決定されると考えられています。その一方で、社会的場面では他者を勘案した意思決定が多く、他者に関する脳内表現の研究も広く行われてきました。その中で、本研究では脳内に表現された他者情報が意思決定プロセスでどのような計算として統合されるのかを研究しました。ここでは、他者への報酬を意思決定に統合する脳計算を解明した成果を述べます。この脳計算は未解明のことが多く、その神経メカニズムは私たち人間の社会性を理解する上で極めて重要であると考えられます。

研究はヒトfMRI実験を計算モデル化解析と組み合わせることで進めました。実験課題は2つの選択肢から一つを選ぶ課題で、それぞれの選択肢には確率的に自己報酬が与えられます。さらには、時折、他者に与えられるボーナス報酬または自己が得られるボーナス報酬が追加されます。この課題設定で計算モデルを適用した解析により、自己報酬に基づく意思決定の脳計算が他者ボーナスによる修飾を受けるプロセスを明らかにしました。

他者ボーナスの提示には、左背外側前頭前野（left dlPFC）に加えて右側頭頂頂接合部（right TPJ）にも脳活動がありました。この左背外側前頭前野は、自己ボーナスの提示でも脳活動がありました。これらは、自己ボーナスと他者ボーナスの間には共通の処理と、さらに他者ボーナスに特有な処理があることを示しています。そして、他者ボーナスの行動選択への影響の大きさは右前島皮質（right AI）の活動が対応し、最終選択に関わる主観的価値は内側前頭前野（mPFC）の活動が対応することを発見しました。その上で、これらの脳活動の関係を調べるコネクティビティ分析を行い、その結果、他者ボーナスについて右側頭頂頂接合部（right TPJ）と左背外側前頭前野（left dlPFC）→右前島皮質（right AI）→内側前頭前野（mPFC）の3段階の脳活動のカスケードがある脳回路を同定しました（図の赤矢印）。一方、自己ボーナスでは、右前島皮質（right AI）を経由せずに、左背外側前頭前野（left dlPFC）→内側前頭前野（mPFC）への直接のカスケードがあることが分かりました（図1 緑矢印）。これらの結果は、他者

の報酬を意思決定に統合するときには、特有の脳回路が働くことを、特に右前島皮質（right AI）が他者の報酬を選択に勘案するための処理を担っていることを示しています。私たちはこの脳回路の特性がさまざまな社会行動の個人差の基盤になっていると推測しています。実際に、社会的価値志向性テストにより、向社会的な人と個人主義的な人に分けただけで、上述の脳活動をさらに詳しく解析すると、向社会的な人と個人主義的な人の中で、異なる脳活動の特性があることも分かりました。このことは、今回発見した回路が、複雑な社会行動の個人差の根底にある可能性を示しています。

本研究の成果により、私たちの社会行動のもとになる神経基盤の一つが特定されました。さらに、この回路の働き方の違いが社会行動の個人差にも関わることが明らかになりました。今後、この神経基盤が他のどのような社会行動の違いに関わるのかなどが定量的に明らかにされていくことが期待されます。

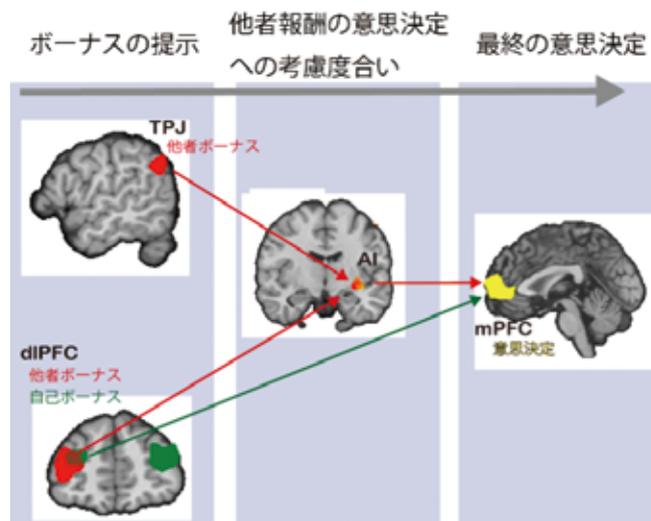


図1 他者の報酬を考慮した意思決定を行うための脳回路

### 主な研究業績：

- 1) Fukuda H, Ma N, et al. (2019) Computing Social Value Conversion in the Human Brain. *Journal of Neuroscience*, 39 (26), 5153-5172.  
<https://doi.org/10.1523/JNEUROSCI.3117-18.2019>

## 予測符号化モデルと自律推論する脳機構との照合

研究代表者：小村 豊（京都大学大学院 人間・環境学研究科 教授）

意思決定は、現在の情報から未来の状況を予測し、複数の選択肢からベストを推論する機能ですが、その過程には、確信度が伴っています。では意思決定と確信度は、同じ認知ソースから生成されているのでしょうか。この問いを明らかにするために、本研究では、動物に wagering 課題を課して、意思決定における判断バイアスと精度の変動にともなって、確信度が、どのように変化するかを、実験データと数理モデルの両面から、検証しました。数理モデルとしては、判断バイアスに即した事前確率を組み込んだ状態で、確信度を計算する現実モデルと、常に中立な事前確率で確信度を計算する反実仮想モデルを用意し、判断バイアスと精度をシミュレートしたところ、反実仮想モデルのみが、実験で得られた wagering データを説明できることが分かりました。この結果は、意思決定と確信度は、少なくとも部分的には異なる認知ソースから生成されていることを示しています。新学術領域で得られた知見を踏まえて、現在は、メタ学習の実態の解明にとりこんでいます。

また自由エネルギー原理の枠組みの中で、実験データを再考することで、確信度に基づく予測推論の機能的意義について、新たな解釈を加える目途がたってきました。

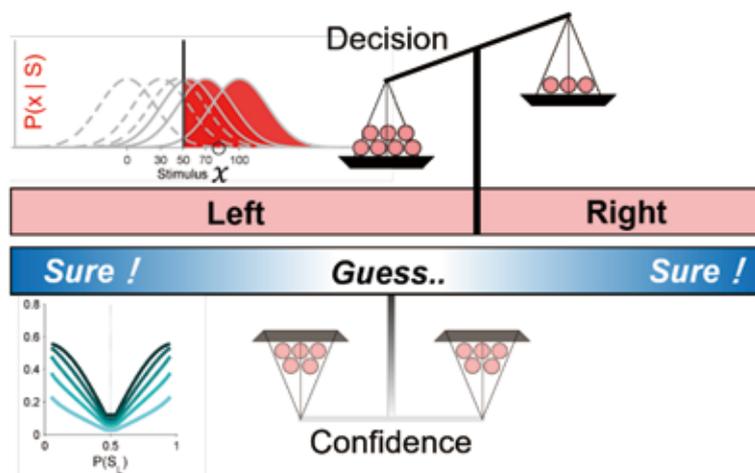


図 予測推論における反実仮想モデル

### 主な研究業績：

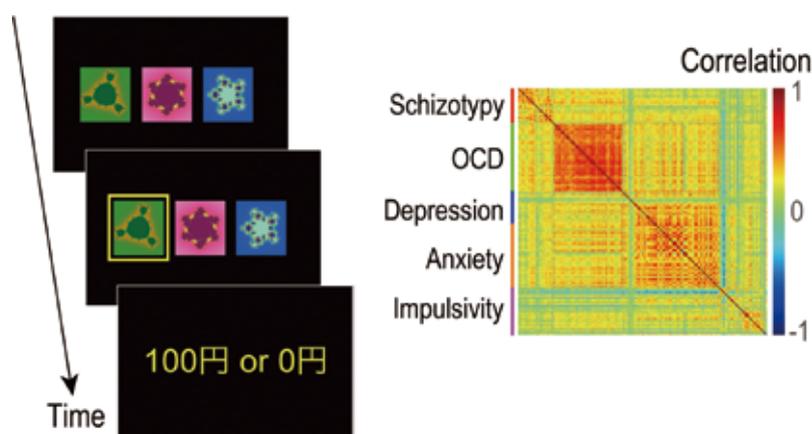
- 1) 知念浩司, 三澤魁旺, 小村豊 (2019). Predictive brains generate our perceptual worlds. 臨床精神医学 48(12) 1399-1402
- 2) 藤本蒼, 野口真生, 小村豊 (2018). メタ認知からみた意識の生物学. 人工知能学会誌 33 (4) 468-471

## 構造学習の脳計算モデル：脳イメージング実験と大規模 WEB 調査による検証

研究代表者：鈴木 真介 (Department of Finance, Faculty of Business and Economics, The University of Melbourne, Associate Professor)

「たくさんの報酬を得る、手に入れた報酬はできるだけ失わない」ための意思決定は我々ヒトの生存に重要な役割を果たします。近年の脳科学の進歩により、これらの意思決定過程は機械学習アルゴリズムの一種である「強化学習」で記述できることが明らかになっています。また、意思決定の障害は抑うつ、不安障害、強迫性障害、統合失調症など様々な精神疾患と関連している可能性が示唆されています。しかしながら、精神疾患関連の症状は相互に複雑に絡み合っているため、「どの疾患・症状が意思決定の障害と関連しているのか?」、「その症状は意思決定のどの過程に影響を与えているのか?」はいまだ不明です。

我々の研究チームは、2000人規模の被験者が参加するオンライン実験を実施し、この問いに取り組みました。その結果、(1) 精神疾患症状はお互いに関連し合っているが、「強迫行為や侵入思考」、「不安や抑うつ」、「その他」といったクラスタに分けられる；(2) その中で「強迫行為や侵入思考」に関連する症状が「報酬や損失からの学習（強化学習モデルの学習率）」を阻害し「同じ行動を繰り返す傾向」を強める、ことを明らかにしました。これらの知見は、人工知能や脳科学の成果を精神疾患の理解に役立てる「計算論的精神医学」の発展に寄与すると考えられます。



図左 オンライン実験における意思決定課題。被験者は図形の選択を繰り返し行い、選んだ図形に応じて確率的に報酬を得る。

図右 精神疾患に関する質問紙データ。各種症状間の相関係数を示している。

### 主な研究業績：

- 1) Suzuki S, Yamashita Y, Katahira K (2021). Psychiatric symptoms influence reward-seeking and loss-avoidance decision-making through common and distinct computational processes. *Psychiatry and clinical neurosciences*, 75, 277-285.  
<https://doi.org/10.1111/pcn.13279>

## 社会的身体表現による個体間相互作用の生成モデル開発

研究代表者：三村 喬生（量子科学技術研究開発機構 脳機能イメージング研究部 研究員）

ヒトや動物は、自らの意志や意図に応じて姿勢や動作のパターンを柔軟に変化させることが知れています。この機能は社会的相互作用において行動表現系を調整する中心的な役割を担うと考えられますが、既存の解析手段では観察に基づく特徴行動（エソグラム）の分類に限られてきました。機械学習を用いることで、観察では分類が難しい微妙な行動表現のパターンを分離・抽出できる可能性があります。

この研究アプローチの実証実験のために、小型霊長類コモン・マーマセットを用いました。この動物は、ヒトと比較可能な神経構造・機能を有し、また高度な社会性認識（公平性や互惠性）を持ちます。マーマセットの行動を死傷なく詳細に記録するために円柱型のケージを用意し、4台の深度カメラをその周囲を囲むように設置しました。合計1020秒間のマーマセットの自由行動の録画データから、モーショントラッキングにより身体の部位と顔の向きの空間座標を算出しました。これらの行動情報に含まれる動作の基本単位を、特殊なノンパラメトリックベイズモデル（HDP-GP-HSMM）により分離・抽出しました（図1）。結果、テスト環境におけるマーマセットの自由運動は3.5秒間の長さの18種類の基本単位に分解できることが示されました。特に、データに含まれる摂餌のタイミングに着目したところ、その前後を含めた3つの行動単位配列が個体を超えて保存された特異的な傾向を持つことが明らかとなりました。従って、この人工知能技術は、動物が「誰で」「どこか」に依らず「何」をしていたのかを、観察者の主観に依らず、データ駆動に推定することに成功しました<sup>1)</sup>。この行動解析技術を用いることで、典型的な行動とそこから逸脱性を高感度に評価できると期待できます。

そこで、神経科学的介入による社会性機能障害モデル動物を作成し、どのような神経機能がどのような社会性行動表現に寄与しているかを調べ

る試みを進めています。これまで2種類の技術を使ったモデル動物の作出に成功しています。1つめは、自閉症におけるリスク要因として知られる胎生期の抗てんかん薬（バルプロ酸）暴露をマーマセットで再現したモデル動物です（図2）。この動物において高解像度拡散テンソル画像法と自動輪郭抽出アルゴリズムにより脳構造を詳細に検討した結果、左右の脳半球を接続する神経繊維の1種（前交連）の減弱などの神経発達上のリスクが見出されました<sup>2)</sup>。現在、行動表現型の解析を進め、家族内での音声コミュニケーションを自然言語処理的アプローチで解析することにより有意な異常をみとめています（発表準備中）。2つめの技術は化学遺伝学です（図3）。これは人工神経受容体（DREADD）を脳の局所の神経細胞にウィルスベクター法で強制的に発現させる手法です。このDREADDに特異的に結合する薬剤を動物に投与することで、一過的に神経細胞の抑制や過興奮を引き起こすことができます。この手法は齧歯類で開発されましたが、我々のチームはMRI、PET、CTなど多様な脳イメージング技術を駆使することで霊長類への応用を進め、マーマセットにおいて世界初となる化学遺伝学の成功事例を報告しました<sup>3)</sup>。更に、社会性を司る神経核の1つである扁桃体にDREADDを発現させることで、薬剤投与後の数時間に限った神経活動の抑制に成功しました。

人工知能技術と神経科学的手法を組み合わせた本研究により、これまで解析が困難だった個体間の複雑な社会的相互作用を定量的に評価する手法と、社会性脳機能の病態モデルを作成する手法の2つが確立しました。これら技術は、複雑な社会性相互作用において神経機能の障害が及ぼす影響を詳細に解析する手段を提供することで、社会性障害の病態の把握や診断の支援において基盤的成果となると期待されます。

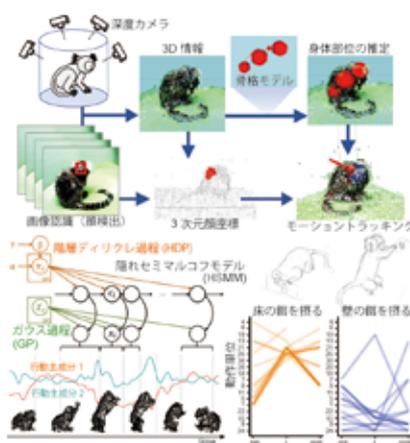


図1 ノンパラメトリックベイズモデル（HDP-GP-HSMM）による行動分節化

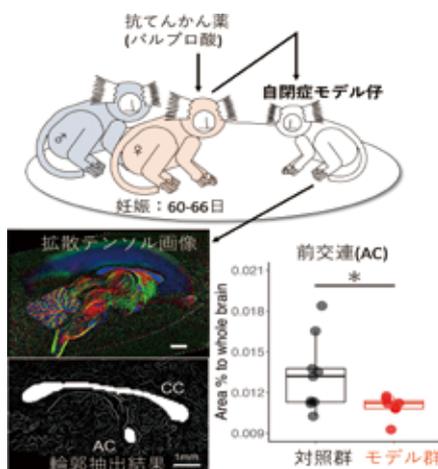


図2 抗てんかん薬（バルプロ酸）の胎生期暴露による自閉症マーマセット

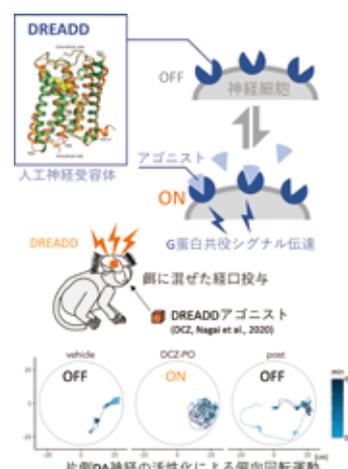


図3 化学遺伝学による神経操作

### 主な研究業績：

- 1) Mimura K (2021) Unsupervised temporal segmentation of marmoset goal-directed behavior, 10<sup>th</sup> International Symposium on Biology of Decision-Making, Paris, France
- 2) Mimura K, Oga T, Sasaki T, Nakagaki K, Sato C, Sumida K, Hoshino K, Saito K, Miyawaki I, Suhara T, Aoki I, Minamimoto T, Ichinohe N (2019). Abnormal axon guidance signals and reduced interhemispheric connection via anterior commissure in neonates of marmoset ASD model, *NeuroImage*, 195, 243-251.
- 3) Mimura K, Nagai Y, Inoue K, Matsumoto J, Hori Y, Sato C, Kimura K, Okauchi T, Hirabayashi T, Nishijo H, Yahata N, Takada M, Suhara T, Higuchi M, Minamimoto T (2021) Chemogenetic activation of nigrostriatal dopamine neurons in freely moving common marmosets, *iScience*, 24, 103066

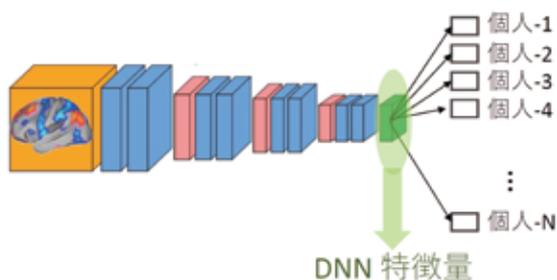
## 深層学習を用いた安静時機能的MRIからの汎用特徴量抽出

研究代表者：山下 祐一（国立精神・神経医療研究センター 室長）

本研究課題では、高解像度MRIデータから階層性・非線形性を伴いながら、自己組織的特徴量抽出が可能な深層ニューラルネットワーク（deep neural networks: DNN）を用いて、脳構造・活動の個人特徴や認知的特性を反映しうる汎用特徴量を抽出する方法の確立を目的としました。開発した汎用特徴量抽出技術を、精神疾患のデータに適用することで、MRIデータから患者個々人の診断・疾患予後・治療反応性予測などを総合的に評価する、精神疾患の評価法のための基盤技術の開発を試みました。具体的には、畳み込みネットワーク（CNN）を用いることで、階層性・非線形性を伴いながら、超高次元3D脳構造画像を0.67%の次元数（15万次元から約1000次元）まで圧縮し、高い精度で再構成することに成功しました<sup>1)</sup>。続いて、開発した高解像度（voxel-based）のMRIデータをDNNで解析する技術を安静時機能的MRI（rsfMRI）に適用することで、脳活動の個人特徴や認知的特性を反映しうる汎用特徴量抽出技術の開発を試みました。開発したCNNに、rsfMRI

の各フレーム（一般的なrsfMRIは10-15分間の撮像時間に対して100-200フレームのMRI画像で構成される）のオリジナル画像（約15万次元）を入力し、出力として“入力画像が誰のものであるか”という個人識別をするよう学習を行いました（図A）。結果として、学習では経験していない未知のサンプル被験者の脳活動（つまり複数のフレーム画像で構成されるrsfMRI画像）を、一人の個人に由来すると認識した上で、学習で経験した人の特徴の組み合わせとして表現できることが確認されました。続いて、開発した個人識別CNNによって抽出した特徴量を用いて、精神疾患（統合失調症）の診断予測を行ったところ、従来の機能的結合解析に基づく特徴量と同等以上の有効性を持つことが確認できました（図B<sup>2)</sup>。本研究で開発した汎用特徴量抽出技術を、統合失調症以外の幅広い精神疾患を対象として適用することで、rsfMRIデータから患者個々人の診断・疾患予後・治療反応性予測などを総合的に評価する、精神疾患の評価法開発に役立つ可能性が期待されます。

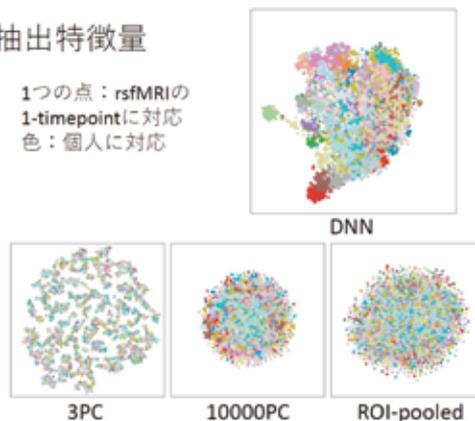
### A. 個人識別CNN



図A rsfMRIの1フレームからそれが誰の脳画像かを識別する3次元畳み込みネットワーク（個人識別CNN）の模式図。出力層の直前のユニットの活性化値を入力から抽出した特徴量として扱う。

### B. 抽出特徴量

- 1つの点：rsfMRIの1-timepointに対応
- 色：個人に対応



図B 抽出特徴量を2次元に圧縮しプロットしたもの。各点がrsfMRIの1フレームに対応し、各点の色は個人に対応する。他の手法と比べて、DNNにより抽出した特徴量は個人ごとにクラスターを形成し、未知のrsfMRIデータから個人の脳活動に対応する特徴量を抽出できていることがわかる。

#### 主な研究業績：

- 1) Yamaguchi H, Hashimoto Y, Sugihara G, Miyata J, Murai T, Takahashi H, Honda M, Hishimoto A, Yamashita Y (2021) Three-dimensional convolutional autoencoder extracts features of structural brain images with a diagnostic label-free approach: Application to schizophrenia datasets. *Front. Neurosci.* 15:652987. doi: 10.3389/fnins.2021.652987
- 2) Hashimoto Y, Ogata Y, Honda M, Yamashita Y (2021) Deep Feature Extraction for Resting-State Functional MRI by Self-Supervised Learning and Application to Schizophrenia Diagnosis. *Front. Neurosci.* 15:696853. doi: 10.3389/fnins.2021.696853

## 自然言語処理と神経生理計測を融合した言語の神経計算モデルの構築と検証

研究代表者：大関 洋平（東京大学 大学院総合文化研究科 講師）

本研究では、記号的生成モデルおよび深層ニューラルネットワークに基づく神経計算モデルを構築し、脳情報処理データで検証することで、自然言語の計算システムを解明することを目的としています。新型コロナウイルス感染症の影響により、脳磁場および皮質脳波による新規の脳情報処理データこそ完成しませんでした。眼球運動による既存の認知処理データを用いて、記号的生成モデルと再帰的ニューラルネットワークを融合した深層生成モデルである再帰的ニューラルネットワーク文法 (Recurrent Neural Network Grammar) が、純粋な再帰的ニューラルネットワークと比べて人間らしいことを示し、更に並列処理による高速化を実現しました。また、アテンション機構を

備えた大規模な深層ニューラルネットワークであるトランスフォーマー (Transformer) が、工学的には世界最高性能を達成しているにも関わらず、強力すぎる作業記憶を持つため人間らしくないことを示し、加えて記号的生成モデルとトランスフォーマーを融合した深層生成モデルであるトランスフォーマー文法 (Transformer Grammar) を開発しています。以上の研究成果は、自然言語処理のトップカンファレンスである ACL および EMNLP にて、3本の論文として発表・出版しています。

なお、本研究課題は、JST さきがけ「信頼される AI の基盤技術」領域「認知・脳情報処理による人間らしい言語処理モデルの開発」へと発展しています。

### 言語：

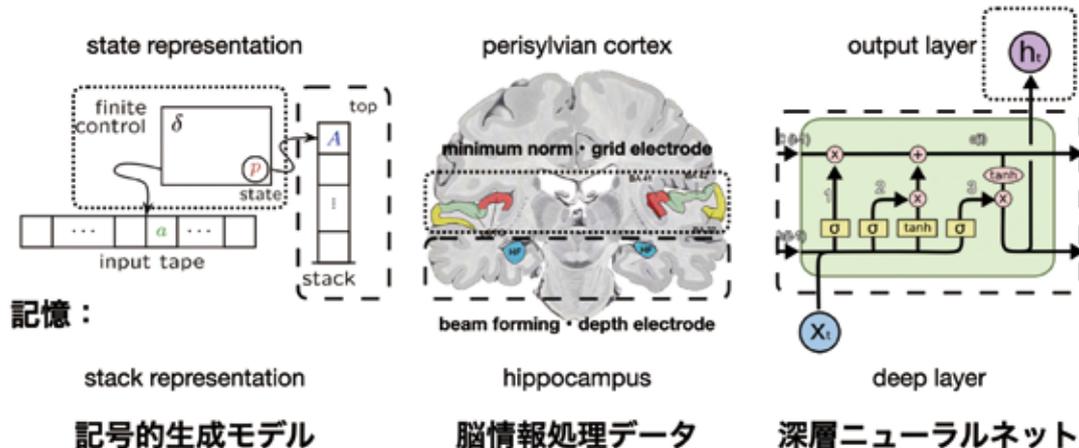


図 神経活動データと神経計算モデルの対照

### 主な研究業績：

- 1) Yoshida R, Noji H, Oseki Y (2021) Modeling Human Sentence Processing with Left-Corner Recurrent Neural Network Grammars. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2964-2973.
- 2) Noji H, Oseki Y (2021). Effective Batching for Recurrent Neural Network Grammars. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL)*, 4340-4352.
- 3) Kuribayashi T, Oseki Y, Ito T, Yoshida R, Asahara M, Inui K (2021). Lower Perplexity is Not Always Human-Like. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL)*, 5203-5217.

# Neural Networks 特集号： Artificial Intelligence and Brain Science

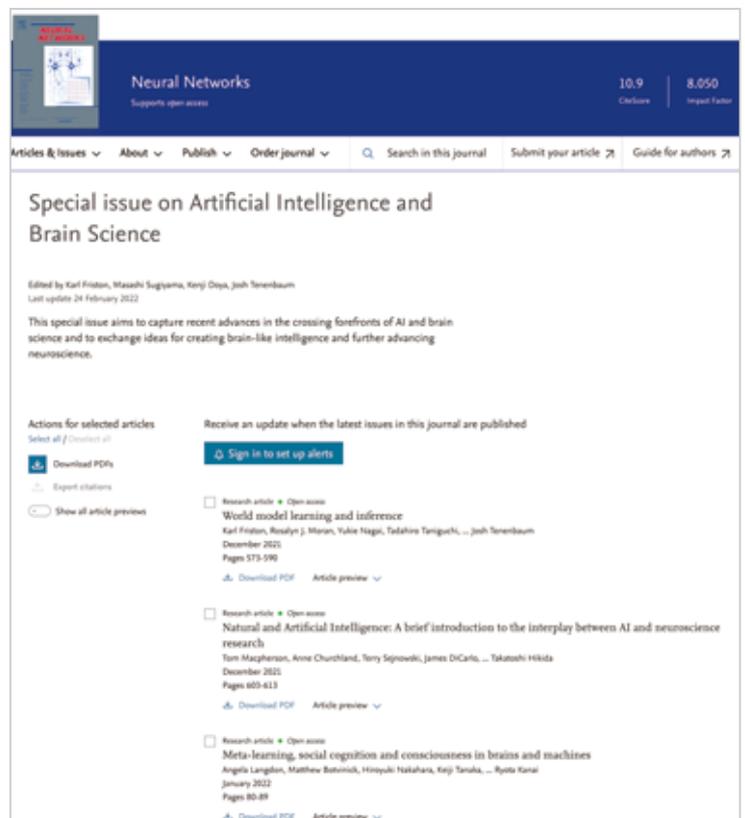
新学術領域「人工知能と脳科学」の主催で2020年10月10～12日にオンライン開催された International Symposium on Artificial Intelligence and Brain Science (<http://www.brain-ai.jp/symposium2020/>)での発表をもとに、国際神経回路学会、日本神経回路学会、欧州神経回路学会の論文誌 Neural Networks (Impact Factor 8.05) の特集号が企画されました。

シンポジウム中およびその後の宣伝により、多数の投稿があり、査読の結果約40本の論文が採択され、Virtual Special Issueとして Elsevier 社の web サイト上に収録されています：

<https://www.sciencedirect.com/journal/neural-networks/special-issue/1017VB2Z2R5>

特集号のコアとなるのは、シンポジウムの各セッションの講演者らの共著による5つの展望論文です。

- Yutaka Matsuo, Yann LeCun, Maneesh Sahani, Doina Precup, David Silver, Masashi Sugiyama, Eiji Uchibe, Jun Morimoto: **Deep learning, reinforcement learning, and world models.**
- Karl Friston, Rosalyn J. Moran, Yukie Nagai, Tadahiro Taniguchi, Hiroaki Gomi, Josh Tenenbaum: **World model learning and inference.**
- Angela Langdon, Matthew Botvinick, Hiroyuki Nakahara, Keiji Tanaka, Masayuki Matsumoto, Ryota Kanai: **Meta-learning, social cognition and consciousness in brains and machines.**
- Tom Macpherson, Anne Churchland, Terry Sejnowski, James DiCarlo, Yukiyasu Kamitani, Hidehiko Takahashi, Takatoshi Hikida: **Natural and artificial intelligence: A brief introduction to the interplay between AI and neuroscience research.**
- Kenji Doya, Arisa Ema, Hiroaki Kitano, Masamichi Sakagami, Stuart Russell: **Social impact and governance of AI and neurotechnologies.**



多彩な講演者による発表の内容をひとつの論文にまとめるのは大変な作業ではありましたが、そのためにシンポジウム後も引き続きオンライン会議やメールにより議論を深めることができたのは大きな収穫でした。人工知能と脳科学の融合研究をリードする研究者らによる成果と構想を集めたこれらの論文が、今後新たな研究を触発することを大いに期待しています。

また一般投稿の論文として、“AI-powered Brain Science” のトピックで約15本、“Brain-inspired AI” のトピックで約20本の論文が収録されています。これらのうち約10本は当領域メンバーや研究協力者によるもので、Perspective論文とあわせ Open Access になっています。

ぜひ Web サイトを訪れダウンロードし、ご一読いただければ幸いです。

(領域代表：銅谷 賢治)

# 新学術領域研究「人工知能と脳科学」メンバーによる 新研究プロジェクトの紹介

## 学術変革領域 B「情動情報解読による人文系学問の再構築」

領域代表者：近添 淳一（株式会社アラヤ 脳事業研究開発室 チームリーダー）

心理学・言語学・経済学・美学といった人文系学問では、人間の行動やその記録から、心的プロセスのモデルを作ります。こうしたモデルを考えるにあたって、情動が人間の行動にどのような影響を及ぼすかを理解することが重要ですが、個人が「どのように感じているか」を直接計測することが難しいことから、情動の働きを直接モデルに取り込むことは簡単ではありませんでした。

機能的MRIは、生きた人間の脳活動を計測できる手法で、被曝などの心配のない安全な手法であることが知られています。近年、機械学習を使った解析技術の進歩によって、脳活動から個人の情動状態を推定することが可能になりつつありますが（Chikazoe et al., 2014; Pham et al., 2021）、解析上の技術的ハードルも高く、脳活動から推定した情動状態の情報を取り込んだ言語学・経済学・美学モデルはほとんどありません。

本研究領域においては、機械学習を用いた機能的MRI研究の専門家である近添（アラヤ）が中心となって、自然言語処理の専門家である持橋大地先生（統計数理研究所）とマイクロ経済学の専門家である渡辺安虎先生（東京大学）、および美学研究の専門家である石津智大先生（関西大学）と協力し、情動から言語・経済・芸術を理解するような新しい学問の枠組みを作っていきます。

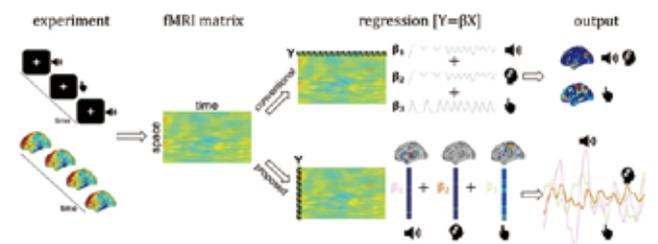


図 情動情報解読のための新規解析手法（小山ら、論文準備中）。情動に関連する脳活動の時系列情報を、情動のテンプレートmapによる回帰で推定する。

## 学術変革領域研究 A「適応回路センサス」計画研究 「環境に応答した行動変容を司る遺伝子・細胞機能ダイナミクス」

計画研究代表者：小林 和人（福島県立医科大学 生体機能研究部門 教授）

学術変革領域研究（A）「神経回路センサスに基づく適応機能の構築と遷移バイオメカニズム、略称:適応回路センサス、領域代表:磯村宜和 東京医科歯科大学教授」では、ヒトや動物の適応行動を担う神経回路の構築や機能的遷移について、1細胞オミクス、細胞と回路の活動計測と活動操作、行動解析、および数理学を統合し、これまでにない高解像度で理解することを目指します。

我々の計画研究班では、ラットが聴覚刺激のピッチの高低に応じて2つのレバーから報酬に連合した正解レバーを選択する刺激弁別学習を行動モデルとした研究をすすめます。これまでに我々は、この刺激弁別学習の獲得プロセスの初期と後期で、線条体の異なる領域の活動が亢進することを見出しました。そこでまず、時期や脳領域に特異的に発現パターンを変える遺伝子をRNAseq法により探索します。同定された遺伝子の機能を、RNA干渉やゲノム編集技術を用いて解析するとともに、各遺伝子を発現する細胞種のプロファイリングをおこないます（分担研究者・松下夏樹 愛知医科大学准教授）。

その行動的機能を、ウイルスベクターを用いた活動操作により解析します（研究分担者・瀬戸川将 獨協医科大学助教）。一連の検討を通じて、適応行動を担う回路動態の分子・細胞メカニズムを解明したいと考えています。

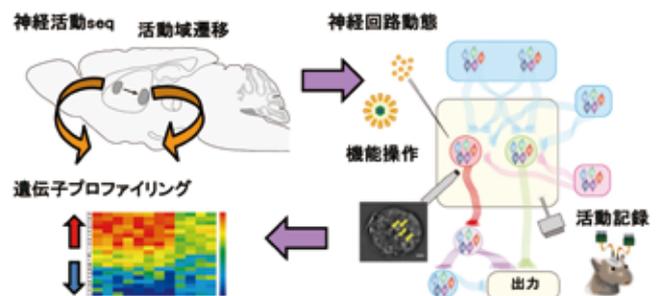


図 変化する環境に適応するための行動の獲得および実行の神経基盤=適応回路センサスを理解するために、大脳基底核を含む神経ネットワークの機能的遷移を司る遺伝子機能・細胞機能・回路動態の解明に挑む。

URL : <https://ac-census.org/>  
 (神経回路センサスに基づく適応機能の構築と遷移バイオメカニズム [略称:適応回路センサス] 公式ウェブサイト)

# JST CREST 「共生インタラクション」領域 「脳表現空間インタラクション技術の創出」

研究代表者：柳澤 琢史（大阪大学高等共創研究院 教授）

ヒトが画像を見ると、脳内では画像の意味を読み解く処理が行なわれます。そこで、様々な画像をヒトが見ている際の脳活動を調べることで、画像の持つ意味や特徴に対応した脳活動の地図を作ることができます。また、ヒトが同じ画像を心の中で想像することでも、同様の脳活動を得ることができ、脳活動の地図を使って、どのような内容を想像したかを推測することができます。我々は、てんかんなどの治療目的で頭の中に電極を入れた患者さんにご協力いただき、様々な画像を見た際の頭蓋内脳波（皮質脳波）を計測し、詳細な地図を作成しています（図1）。特に、画像の持つ意味や特徴を表すために、AI技術を使うことで脳活動と対応付いた地図を作ることができました。さらに、この地図に基づいて、ヒトが想像した画像を画面に出力する技術を開発しています。この技術はヒトの想像や思考内容に対応した多様な脳活動と外界の多様な意味に応じた計算機内の表現空間とを結びつける新しいインタラクション技術です。本研究領域では、この技術の神経科学的基

盤解明と臨床応用を目指して研究をしています。なお、この研究は新学術人工知能と脳科学での研究課題から発展しました。



図1 CREST 研究のイメージ図  
脳内の意味空間と外界の視覚情報空間との間をAIでつなげることで、ヒトと脳内表現空間とのインタラクションを実現する。

プロジェクトHP：

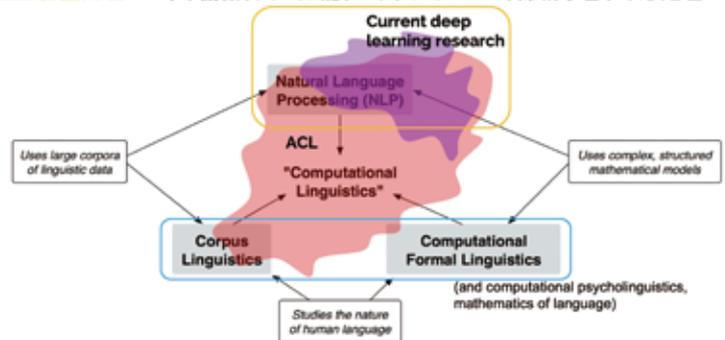
<https://www.med.osaka-u.ac.jp/pub/nsurg/yanagisawa/CREST/>

# JST さきがけ 「信頼されるAIの基盤技術」領域 「認知・脳情報処理による人間らしい言語処理モデルの開発」

研究代表者：大関 洋平（東京大学 大学院総合文化研究科 講師）

この度、科学技術振興機構 さきがけ「信頼されるAIの基盤技術」領域において、研究課題「認知・脳情報処理による人間らしい言語処理モデルの開発」が採択されました。文部科学省 新学術領域研究「人工知能と脳科学の対照と融合」領域における研究課題では、自然言語の計算システムを解明するという科学的な目的（工学→科学）を中心に据えていましたが、本研究では、人間らしい言語処理モデルの開発という工学的な目的（科学→工学）に方向性をシフトしています。また、新学術領域研究で成果が得られた認知処理の研究を継承しつつ、新型コロナウイルス感染症の影響により完成しなかった脳情報処理の研究へと展開させたいと考えています。特に、自然知能研究としての言語の認知・脳科学と人工知能研究としての自然言語処理を融合することで、認知・脳情報処理に学んだ、人間らしい言語処理モデルを開発することを目的とします。具体的には、クラウドソーシングと脳機能イメージングで計測される人間の言語処理と、記号的生成モデルと深層ニューラルネットワークで実装される機械の言語処理を対照する過程で、言語処理のリバースエンジニアリングを目指します。

自然言語処理：工学、機械、人工知能、コネクショニズム、帰納学習、即応処理



認知・脳科学：科学、人間、自然知能、シンボリズム、演繹学習、熟考処理

図 言語の認知・脳科学と自然言語処理の融合

# イベント情報

共催、協賛、後援、関連イベント

## Neuro2022 in 沖縄

日程：2022.6.30-7.3

場所：沖縄コンベンションセンター、宜野湾市立体育館、ラグナガーデンホテル（on-site/online ハイブリッド）

Web: <https://neuro2022.jnss.org>

日本神経科学学会、神経化学会、神経回路学会の合同大会「Neuro2022」が2022年6月30日から7月3にかけて、沖縄コンベンションセンターで開催されます。「人工知能と脳科学」からも領域代表の銅谷賢治が神経科学学会の大会長、計画研究メンバーの谷口忠大が神経回路学会のプログラム委員長としてその企画をリードしています。

神経科学大会が沖縄で開催されるのは初めてで、多くの方々から関心を寄せられています。新型コロナの状況が見通せない中、オンサイト/オンラインのハイブリッド開催を予定していますが、1月に締め切ったアブストラクトは昨年、一昨年に大きく上回る投稿数に達しています。

プレナリー講演には、意思決定の神経回路とマウス全脳の光学活動計測で注目を集めるAnne Churchland博士（UCLA）、神経再生機構の研究で著名なMartin Schwab博士（University of Zurich）、シナプス可塑性のトランスクリプトーム、メタボローム研究をリードするErin M. Schuman博士（Max Planck Institute）を招いています。その他、Brain Prizeなど受賞者による記念講演、特別講演、教育講演、40以上のシンポジウムなど多彩なプログラムを準備しています。事前登録は5月11日まで。ぜひ多くの皆様に沖縄でお目にかかることを楽しみにしております。



## The 2nd International Symposium on Artificial Intelligence and Brain Science

2020年にオンライン開催し好評を博した当領域の国際シンポジウム（<http://www.brain-ai.jp/symposium2020/>）の続編を、Neuro2022のサテライトイベントとしてとして以下のとおり開催します。

日程：2022.7.4-5

場所：沖縄科学技術大学院大学（<https://groups.oist.jp/conference-venues>）（on-site/online ハイブリッド）

招待講演者：Matthew Botvinick (DeepMind), Karl Friston (UCL), Ryota Kanai (ARAYA), Maneesh Sahani (UCL), Terrence Sejnowski (Salk Institute), Masashi Sugiyama (RIKEN AIP), Jun Tani (OIST), Naotsugu Tsuchiya (Monash University), Xiao-Jing Wang (NYU)

Web: <http://www.brain-ai.jp/symposium2022/>

今後さらに講演者を加え、ポスター発表も募集する予定です。今後詳細は領域webサイト（<http://www.brain-ai.jp/>）他にてアナウンス予定です。ぜひこちらも大勢のご参加をお待ちしております。



発行 / 編集 新学術領域研究「人工知能と脳科学の対照と融合」  
お問い合わせ 新学術領域研究「人工知能と脳科学の対照と融合」事務局  
Mail [ncus@oist.jp](mailto:ncus@oist.jp)  
2022年3月発行

[www.brain-ai.jp](http://www.brain-ai.jp)