

2023 年度年次報告書

信頼される AI システムを支える基盤技術

2021 年度採択研究代表者

竹内 一郎

名古屋大学 大学院工学研究科  
教授

AI 駆動仮説の静的・動的信頼性保証と医療への展開

主たる共同研究者：

大島 孝一（久留米大学 医学部医学科 教授）

川口 淳（佐賀大学 大学院医学系研究科 教授）

佐久間 淳（東京工業大学 情報理工学院 教授）

松井 茂之（名古屋大学 大学院医学系研究科 教授）

## 研究成果の概要

生成 AI の出現により科学研究や技術開発のプロセスが変革されつつある。研究開発対象に関するデータに AI を適用することで科学研究や技術開発に関する新しい仮説が自動的に提案されるようになっており、通常の知識や経験では考えられない有望な仮説を生み出す可能性がある。しかしながら、AI はノイズからでも説得力のある仮説を生成できてしまうため、従来の統計的検定では AI によって生成された仮説を適切に評価できず、新たなアプローチが必要である。AI によって生成された仮説の統計的信頼性を正しく評価するためには、その仮説が AI によって生成されたことを踏まえたうえで評価するしくみが必要となる。本課題では選択的推論法に焦点を当ててこれまで研究を進めてきたが、この推論法は AI によって生成されたことを条件付けたうえで仮説を評価できる数理的な基盤を提供しており、生成 AI 時代の統計的検定と位置づけられる。

本課題では、これまで、選択的推論法に関する方法開発と医療分野における実証にとりくんできた。方法開発では、選択的推論法の検出力向上と適用範囲拡大に取り組んできた。2023 年度は、方法開発において、深層学習に対する選択的推論の枠組を構築したことが最も重要な成果となる。前年度までに特定の構造の深層学習モデルに対する選択的推論法を構築できていたが、これは異なる構造の深層学習モデルのごとに選択的推論のコアとなる部分を個別実装する必要があった。具体的には、選択的推論では AI(深層学習モデル)が特定の仮説を選択するイベントを数理的に特徴づけてソフトウェア実装する必要があった。2023 年度に行った研究により、通常の深層学習の実装と同一の実装コストで特定のクラスの深層学習モデルに対する選択的推論が可能となった。このしくみを活用することにより、深層学習の主要な説明指標である顕著性マップ (saliency map) の統計的信頼性を適切に定量化することを可能にした(主な研究成果 1)。また、佐久間 G では予測結果の説明にとって重要な特徴を選択する方法に対し、統計的信頼性を定量化することに成功した。さらに、松井 G では、個別化医療における適応的な臨床試験において適切な統計的評価法を確立した。これらの方法開発の成果は、いずれもアルゴリズムの複雑さ故に統計的有意性の定量化が困難であった問題であり、AI を用いた科学研究・技術開発における信頼性評価における顕著な成果であると言える。一方、医療実証では悪性リンパ腫の病理診断の課題に取り組んできた。三好 G の研究室(久留米大学病理教室)では、年間数千例の悪性リンパ腫の病理診断を行っており、これは国内では類を見ず、国際的に見ても最大クラスのものである。従来の病理診断は、病理断片が添付されたプレートを顕微鏡で観察することで行われていたが、本課題では、デジタル化したうえで、AI によるサポートのもとで病理診断を行うプロトタイプを作成を目指している。これまで、約 5000 例の症例のデジタル化を完了し、サブタイプ分類、腫瘍細胞同定、類似症例検索などの AI 技術の開発を行ってきた。2023 年度は、特に、病理診断特有の類似症例検索の方法を開発し、その成果は病理画像分析のトップジャーナルに採択された(主な研究成果 2)。また、川口 G では、脳画像分析の新たな方法を開発し、異なる部位の関連を同定することが可能となった。これらの成果は医療画像解析として顕著な成果であるが、今後はこれらの医療画像解析に対して選択的推論などを用いた信頼性定量化を進める予定である。

竹内グループでは、主に、深層学習モデルにおける選択的推論法の一般化を行った。昨年度の成果を発展させることで、特定のクラスの構造を持った深層学習モデルの選択的推論が、追加の実装コストなく実施できるように拡張した。この成果により、深層学習の説明指標として様々な場面で活用されている CAM や Grad-CAM に対する統計的信頼性評価を可能にした。他にも、セグメンテーション、変化点検知などのアルゴリズムに対して選択的推論法を構築した。

佐久間グループは、深層学習モデルの敵対的環境下での攻撃と防御について研究し、ユーザが指定したモデルでのみ正しく機能する可逆的敵対的事例や、未知の変化に対応する防御手法を開発した。また、ブラックボックス分類器に対する抽象概念を用いた説明可能性の研究を行い、特定の概念が分類結果に因果的影響を与える学習プロセスを提案した。これにより、深層学習の説明性とロバスト性を向上させることが期待される。

松井グループは個別化医療を中心に、統計と機械学習を用いた幅広い研究を展開している。主な成果には、臨床試験の適応的デザインに関する新統計手法の開発、データ駆動仮説の効率的な探索フレームワーク構築、脳画像解析や少数症例を用いた安全性評価の手法が含まれる。さらに、治療選択の支援ツール開発や多くの疾患領域での診断法の共同研究を行い、AI 診断の国際規制ガイドラインに基づいた評価基準の整理も進めている。

川口グループは脳画像 AI 解析の信頼性を調査し、アルツハイマー病の脳画像データや COVID-19 の胸部画像データを使用して AI モデルの比較と信頼性評価を行っている、脳画像の前処理方法や心房細動患者の左房平均電位を予測する機械学習モデルを開発し、これらの成果は複数の学会で発表された。さらに、脳画像の関連解析の数理的な研究の成果が同分野の論文誌に採択された。

三好グループは、AI 病理診断システムの構築に向けて悪性リンパ腫症例のデジタル化を推進し、類似症例検索のシステム開発を行った。この方法では、病理画像の特徴に起因する問題をマルチプルインスタンス学習のアプローチで解決する新たな類似症例検索の枠組を構築した。また、悪性リンパ腫の細胞の多様性を分析する AI ツールを開発し、医療分野の論文誌に採択された。

#### 【代表的な原著論文情報】

- 1) Miwa D., Duy V.N.L., Takeuchi I. Valid P-Value for Deep Learning-driven Salient Region. Proceedings of The International Conference on Learning Representation (2023)
- 2) Hashimoto N., Takagi Y., Masuda H., Miyoshi H., Kohno K., Nagaishi M., Sato K., Takeuchi M., Furuta T., Kawamoto K., Yamada K., Moritsubo M., Inoue K., Shimasaki Y., Ogura Y., Imamoto T., Mishina T., Tanaka K., Kawaguchi Y., Nakamura S., Ohshima K., Hontani H., Takeuchi I. Case-based Similar Image Retrieval for Weakly Annotated Large Histopathological Images of Malignant Lymphoma Using Deep Metric Learning. Medical Image Analysis: vol.85, 102752 (2023)

- 3) Das D., Ndiaye E. and Takeuchi I. A Confidence Machine for Sparse High-Order Interaction Model. Stat:e633 (2024)
- 4) Mori J., Furukawa R., Teranishi I., Sakuma J. Heterogeneous Domain Adaptation with Positive and Unlabeled Data. Proceedings of 2023 IEEE International Conference on Big Data (2023)
- 5) Kawaguchi A. Network-based diagnostic probability estimation from resting-state functional magnetic resonance imaging. Mathematical biosciences and engineering: 20(10)p17702-25 (2023)