

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

2022 年度
年次報告書

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 (筑波大学 システム情報系 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

本研究の目的は、AIによって得られる仮説の統計的信頼性を定量化する数理情報技術を構築し、その有効性を医療分野において実証することである。目的を達成するため、本チームは機械学習を専門とする2つの研究グループ(竹内 G, 佐久間 G), 生物統計を専門とする2つの研究グループ(松井 G, 川口 G), 病理診断を専門とする1つの研究グループ(大島 G)から構成されている。機械学習グループでは、主に、深層学習モデルに対する統計的有意性評価を行う方法の開発などを行った。深層学習のように複雑な演算を行う場合、 p 値や信頼区間などの統計的有意性指標を正しく求めることは困難であったが、本研究では、選択的推論(Selective Inference)と呼ばれる技術を拡張することで、理論的に妥当性の保証された p 値や信頼区間を求めることができるようになった。本研究成果は機械学習分野の最難関国際会議である Neural Information Processing Systems (NeurIPS) 2022 に採択されるなど、国内外から高い評価を受けている(代表的な原著論文情報 1-4)。生物統計グループでは、主に、個別化医療と脳画像解析の課題において統計的有意性を定量化するとりくみを実施している。個別化医療では、一部の症例に特異的に見られる異常や一部の部位に特異的に見られる異常を同定することが必用である、データに基づいて症例や部位の選択が行われるため、統計的有意性を正しく評価するためには、選択バイアスの補正が必用となる。本研究では、これらの選択バイアスの補正を選択的推論によって実現する方法の開発にとりくんでいる。病理診断グループでは、主に、悪性リンパ腫の症例データベースの構築と病理診断システム構築のための要素技術の開発を行った。まず、悪性リンパ腫の症例データベースとしては国内最大となる約 5000 症例のデータベースを構築した。本データベースの症例において、類似症例検索システム、亜型分類システムなどの予想技術の開発を行い、病理診断医による性能主観評価実験などを行った。これらの要素技術の研究成果は、医療画像分野の専門誌(代表的な原著論文情報 5)に採択された。

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構築と実証を行った。従来の選択的推論は線形モデルの特徴選択アルゴリズムなど、単純なアルゴリズムに対してのみ適用可能であったが、本研究では適用条件を大幅に緩和することに成功し、深層学習モデルなどの複雑なアルゴリズムへの選択的推論の適用を可能とした。具体的には、従来の選択推論は、選択イベントが線形不等式や二次不等式の積集合で特徴づけられる場合に適用可能であったが、本研究では、この適用条件を大幅に緩和し、線形不等式や二次不等式の積集合の和集合で特徴づけられる場合でも適用可能とした。これにより、多くのクラスの畳込みニューラルネットワークへ選択的推論が適用できるようになった。本研究では、深層学習における統計的推測問題として、特に、医療画像のセグメンテーション結果、ならびに、顕著性マップに対する統計的信頼性の問題を考え、理論的な妥当性の保証された p 値や信頼区間を付与できる方法を開発した。本研究は深層学習の結果に対して、有限サンプルでも妥当性の保証された統計的有意性指標を与える初めての結果であり、深層学習を医療分野などの意思決定に利用する場合に有益である。実際、医療画像へ適用を行ったところ、選択イベントを考慮しない従来の方法では、誤検出が高くなってしまいうるに対し、提案法では誤検出率を有意水準(例えば 5%)未満に正しく制御することができるようにな

った。

佐久間グループでは主に(A)深層学習モデルに対する変動/敵対的環境下における攻撃とそれに対するロバスト性および(B)深層学習モデルに対する説明性に関する研究を行った。(A)の攻撃/防御研究においては(1)ユーザが指定したAIモデルでは画像を正しく認識・利用できるが、それ以外のAIモデルでは敵対的事例として認識され、正しく認識できない可逆的敵対的事例(2)深層学習を使ったコンテンツベース画像検索において、敵対的事例に対する理論保証付き防御手法、(3)未知の環境変化に対応する敵対的ドメイン汎化手法等の研究を通じて、攻撃者/不確実性存在下でもロバストに動作する深層学習モデルを実現する方法論を構築した。(B)の説明性研究においては、「データXはA、Bを持ち、Cを持たないので、クラスYに分類される」という形式のブラックボックスな深層学習分類器に対する抽象概念を用いた説明を可能にする研究を行った。説明に適した概念を発見できる構造的生成モデルを提案し、特定の概念が分類器の出力に大きな因果的影響を与えるように促す学習プロセスを提案した。これを発展させることで、悪性リンパ腫の分類結果に抽象概念による説明を付与できると期待できる。

松井グループでは、個別化医療を中心に幅広い研究テーマについて統計・機械学習の手法とその実践に関する研究を行った。データ駆動仮説の検証との関係では、治療開発の臨床試験の適応的デザインと関連して、標本サイズの再設計、患者集団選択後の治療効果の推定(選択的推論の一種)について新たな統計的手法を開発した。併せて、データ駆動仮説の効率的な探索と検証のための自己対照デザインに基づくデータ解析の枠組みを構築した。他にも、脳画像データを用いた疾患関連解析における階層・潜在構造モデリングと多重検定、少数症例を用いた安全性評価における能動的機械学習などについて手法の開発研究を行った。統計・機械学習手法の実践研究では、医療現場での治療選択の意思決定支援ツール開発の一環として、治療前の患者背景変数、治療後の中間変数を用いた動的予後診断システムを開発した。また、救急、消化器内科、呼吸器内科、産婦人科などの疾患領域において、診断法開発と評価に関する多くの共同研究を行った。以上の検討とは別に、AI診断に関する欧米の規制ガイドライン、検証的臨床試験の計画と報告に関するガイドライン等を調査し、今後の包括的な信頼性評価のためのAI診断の評価・検証規準について整理した。

川口グループでは、脳画像AI解析における信頼性について文献調査を行った。そして、実際にアルツハイマー病脳画像データに対して、AI解析を行いモデルの比較、信頼性評価方法について検討し、これらの内容について学会発表の準備を行った。さらには、COVID-19胸部画像データにおいて診断に有用なAIモデルを検討し、学会発表を行った。選択的推測のために、脳画像データの前処理について、方法及びソフトウェア操作方法について検討を行い、実データを整備し提供した。循環器領域の研究として、心房細動患者の左房平均電位を術前に(外来で)予測するために機械学習法を用いたモデル式の検討を行って、学会発表を行った。無作為2群間比較の臨床試験において、クラスタリング手法を用いた解析を行い、論文投稿準備を進めた。

大島グループでは、AI 病理診断システムの構築を目的とし、悪性リンパ腫症例の収集および整備を行った。現在 2018 および 2019 年の診断症例のデジタル化および症例データの整備が完了しており、合計で約 5000 症例のデータが利用可能となっている。またアルゴリズムの開発に関しても、病理標本画像と患者情報を組み合わせたマルチモーダルクラス分類の開発を行った。既存のデジタル病理の研究では主に画像のみを対象としているが、実際の病理診断では患者情報を参照しながら診断を行うことから、開発予定の AI 病理診断システムにおいても画像と患者情報を組み合わせた統合的な解析が重要になると考えられる。本グループで開発した分類手法では、トランスフォーマー機構を用いたマルチモーダル分類により患者情報がどのように分類結果に寄与するかを示す説明性の高いモデルを実現した。(研究業績 5) マルチモーダル分類として画像とフローサイトメトリー(FCM)を扱った研究も行った。提案手法ではリンパ腫の持つ階層構造のクラス関係を利用し、画像と FCM を上手く組み合わせることで従来手法よりも高い分類精度および説明性を示した。この研究は 2022 年度 AIP チャレンジ「リンパ腫病理におけるマルチモーダル病型分類」の内容となり、現在論文投稿準備中である。

【代表的な原著論文情報】

- 1) V.N.L. Duy et al. "Quantifying Statistical Significance of Neural Network-based Image Segmentation by Selective Inference." *Advances in Neural Information Processing Systems (NeurIPS)*, Dec. 2022.
- 2) V.N.L. Duy et al. "More Powerful Conditional Selective Inference for Generalized Lasso by Parametric Programming." *Journal of Machine Learning Research*, Nov. 2022.
- 3) H. Kato et al. "Safe RuleFit: Learning Optimal Sparse Rule Model by Meta Safe Screening." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Feb. 2023.
- 4) E Ndiaye et al. "Root-finding Approaches for Computing Conformal Prediction Set." *Machine Learning*, Nov. 2022.
- 5) Y. Takagi, et al. "Transformer-based personalized attention mechanism for medical images with clinical records." *Journal of Pathology Informatics*, 14, 100185, Jan. 2023.