

未来社会創造事業（探索加速型）

「超スマート社会の実現」領域

終了報告書（本格研究）

令和 2 年度採択研究開発代表者

[研究開発代表者名:石川 冬樹]

[大学利用機関法人 情報・システム研究機構 国立情報学研究所 アーキテクチャ科
学研究系・准教授]

[研究開発課題名:機械学習を用いたシステムの高品質化・実用化を加速する
"Engineerable AI"技術の開発]

実施期間 : 令和 3 年 1 月 1 日～令和 7 年 3 月 31 日

§1. 研究実施体制

(1)「NII」グループ(国立情報学研究所)

- ① 研究開発代表者:石川 冬樹 (国立情報学研究所 アーキテクチャ科学研究系、准教授)
- ② 研究項目
 - ・ 深層学習修正技術の検討
 - ・ 自動運転における認知 AI への深層学習調整・修正技術の応用
 - ・ 自動運転における認知 AI の安全性ベンチマークの策定

(2)「科学大」グループ(東京科学大学)

- ① 研究開発代表者:鈴木 賢治 (東京科学大学 総合研究院 未来産業技術研究所、教授)
- ② 研究項目
 - ・ 少数データ学習可能な深層学習モデルの検討
 - ・ 少数データモデルの診断支援 AI への応用

(3)「北大」グループ(北海道大学)

- ① 研究開発代表者:松原 崇 (北海道大学 大学院情報科学研究院、教授)
- ② 研究項目
 - ・ 構造を扱う深層学習モデルの検討

(4)「九大」グループ(九州大学)

- ① 研究開発代表者:趙 建軍(九州大学 システム情報科学研究院、教授)
- ② 研究項目
 - ・ 深層学習修正技術の検討
 - ・ 自動運転における認知 AI への深層学習修正技術の応用

(5)「東大」グループ(東京大学)

- ① 研究開発代表者:馬 雷(東京大学 情報理工学系研究科 情報科学科、准教授)
- ② 研究項目
 - ・ 深層学習修正技術の検討
 - ・ 自動運転における認知 AI への深層学習修正技術の応用

(6)「早大」グループ(早稲田大学)

- ① 研究開発代表者:鷺崎 弘宜 (早稲田大学 基幹理工学部情報理工学科、教授)
- ② 研究項目

- ・ 高信頼な機械学習応用システムのためのモデリング・分析技術の検討
- ・ 高信頼な機械学習応用システムのためのフレームワークの検討

(7)「国立がん研究センター」グループ(国立がん研究センター)

- ① 研究開発代表者:三宅 基隆 (国立がん研究センター中央病院 放射線診断科、医長)
- ② 研究項目
 - ・ 医療診断 AI のためのデータ収集・整備

(8)「慶大」グループ(慶應義塾大学)

- ① 研究開発代表者:陣崎 雅弘 (慶應義塾大学 医学部放射線科(診断)、教授)
- ② 研究項目
 - ・ 医療診断 AI のためのデータ収集・整備

§2. 研究開発成果の概要

機械学習技術、特に深層学習技術を用いた AI の実用化が盛んに追求されているが、その品質、特に安全性や信頼性に対して大きな懸念がある。本研究開発課題では“**Engineerable AI (eAI)**”というビジョンを提唱し、AI システムの開発・品質保証・運用を安定化・効率化することで、その高度化および実用化を大きく加速させることを目標としている。具体的には、大量データへの依存性および性能の制御困難性によりドメイン・システム固有の要件への対応が難しいという課題に対し、AI 構築のための知識融合型深層学習技術、AI 性能の調整・修正のための深層学習自動デバッグ技術の二つを軸とした研究開発を行った。さらに、医療・交通の二分野における重要な課題に取り組むことにより、これら技術の POC (Proof-of-Concept、概念実証)を行った。医療 POC においては、非専門医や従来 AI にとって検出が難しい病変に強い診断支援について、交通 POC においては、膨大な状況を扱う先端運転機能に対する安全性論証について、それぞれ実証に取り組んだ。以上の成果として、まず、要求やリスクを分析し、AI の構築や更新を適切に行うための技術群を含むフレームワークが得られた。医療 POC においては、2大がんの希な非典型病変の検出において、従来 AI の見落としを 50%以上削減することができた。交通 POC においては、運転シーンにおける異なる種別の誤分類を評価する 16 個の指標について、必要な制約を満たしつつ半数以上の指標を向上する性能の調整・修正能力を示すことで、安全性に対するアライメントプロセスを実証することができた。以上の成果により、対象ドメイン・システムの要求・リスクを系統的・包括的に分析し、それに対応する個々の「信頼できる AI システム」を仕立て上げる工学技術の基盤を確立することができた。

【代表的な原著論文情報】

1. Preedan W., Suzuki K., Kondo T., Kobayashi M., Tanaka H., Ishioka J., Matsuoka Y., Fujii Y., and Kumazawa I.: Urinary Stones Segmentation in Abdominal X-Ray Images Using Cascaded U-Net Pipeline with Stone-Embedding Augmentation and Lesion-Size Reweighting Approach. IEEE Access 11: 25702-25712, 2023
2. Davide Li Calsi, Matias Duran, Xiao-Yi Zhang, Paolo Arcaini, Fuyuki Ishikawa, Distributed Repair of Deep Neural Networks, The 16th IEEE International

Conference on Software Testing, Verification and Validation (ICST 2023), pp. 83-94, April 2023

3. Hua Qi, Zhijie Wang, Qing Guo, Jianlang Chen, Felix Juefei-Xu, Fuyuan Zhang, Lei Ma, Jianjun Zhao, ArchRepair: Block-Level Architecture-Oriented Repairing for Deep Neural Networks, ACM Journal on Transactions on Software Engineering and Methodology (TOSEM), Vol.32 No.5, pp.1-31, 2023
4. Jati H. Husen, Hironori Washizaki, Jomphon Runpakprakun, Nobukazu Yoshioka, Hnin Thandar Tun, Yoshiaki Fukazawa, Hironori Takeuchi, “Integrated Multi-view Modeling for Reliable Machine Learning-Intensive Software Engineering,” Software Quality Journal, Vol. 32, pp. 1239–1285, Springer-Nature, 2024.
5. Hironori Takeuchi, Jati H. Husenb, Hnin Thandar Tun, Hironori Washizaki, Nobukazu Yoshioka, “Enterprise Architecture-based Metamodel for Machine Learning Projects and its Management,” Future Generation Computer Systems, Elsevier, Vol. 161, pp. 135-145, 2024.