＜年次報告書＞
第 1 年度（2025 年度）
日 ASEAN 科学技術・イノベーション協働連携事業（NEXUS）
日本－シンガポール「AI」研究領域

研究課題名（和文）	統計的に異質なエッジクラウドネットワークにおける効率的かつプライバシー保護された大規模マルチモーダルモデルの学習と推論
研究課題名（英文）	Efficient and Private Large Multi-Modal Model Training and Inference over Heterogeneous Edge-Cloud Networks
研究期間	2025 年 4 月 1 日～2028 年 3 月 31 日

氏名	所属 部署　役職
曹 洋	東京科学大学 情報理工学院　准教授
蘇 春華	会津大学 コンピュータ理工学部　上級准教授
RASHED Essam	兵庫県立大学 大学院情報科学研究科　教授

1. 研究進捗状況と成果

本年度は、エッジ・クラウド環境における大規模マルチモーダルモデル（LMM）の効率的かつプライバシー保護された学習・推論基盤の確立を目的とし、連合学習および安全推論に関する要素技術の設計・実装・評価を中心に研究を推進した。

まず、統計的に異質なデータ分布を有する連合学習環境において、従来の重み平均型集約では十分な汎化性能が得られないという課題に対し、クライアントが生成したクラス別プロトタイプを活用する Multi-domain Prototype-based Federated Fine-Tuning（MPFT）を提案した。本手法は、生データを共有せずにサーバ側でアダプタを最適化することで、in-domain および out-of-domain 双方の性能向上を実現し、さらに通信ラウンドを削減する軽量な学習枠組みを構築した。本成果は ICLR 2025 に採択された。

次に、エッジ端末上での安全な大規模モデル推論に関して、TEE（Trusted Execution Environment）を活用した推論フレームワーク AegisGuard を開発した。従来の全アダプタ TEE 保護方式では通信遅延が大きいという問題に対し、強化学習に基づく層別感度推定により保護すべきアダプタのみを選択的に TEE 内へ配置し、さらに構造的圧縮を行うことで、モデル盗用耐性を維持しつつ推論遅延を大幅に削減することを実証した。本成果は NeurIPS 2025 に採択された。

また、本年度は国際研究交流の促進にも重点を置き、日本・シンガポール研究チーム間で定期的なオンラインミーティングおよび技術討論を実施した。研究課題の整理、実験設計、論文化作業を共同で進めることで、連合学習と安全推論を統合する研究基盤を構築した。これらの継続的な交流活動により、国際共同研究体制が強化され、上記国際会議での成果発表へと結実した。

2. 代表的な成果事例

1. Jingyuan Zhang, Yiyang Duan, Shuaicheng Niu, Yang Cao, Wei Yang Bryan Lim. Enhancing Federated Domain Adaptation with Multi-Domain Prototype-Based Federated Fine-Tuning (MPFT). In International Conference on Learning Representations (ICLR), 2025.
2. Che Wang, Ziqi Zhang, Yinggui Wang, Tiantong Wang, Yurong Hao, Jianbo Gao, Tao Wei, Yang Cao, Zhong Chen, Wei Yang Bryan Lim. AegisGuard: RL-Guided Adapter Tuning for TEE-Based Efficient & Secure On-Device Inference. In Advances in Neural Information Processing Systems (NeurIPS), 2025.