

先端国際共同研究推進事業

2023 年度採択

次世代のための ASPIRE

半導体分野

2023 年～2024 年度

年次報告書・公開版

研究課題名	A S I C 設計ユニバーサル化に向けた、粗粒度ロジックア レープラットフォームの創生
日本側研究代表者	小菅 敦丈 東京大学大学院 工学系研究科 講師
相手側研究代表者	Priyanka Raina, Assistant Professor, Stanford University
研究期間	2024 年 2 月 1 日～2027 年 3 月 31 日

1. 研究成果の概要

① 研究構想にかかる成果

<実施したこと>

2024 年よりスタンフォード大学に日本側の若手研究者が滞在し、密な議論をかわしながら AI プロセッサに関する実装・検証を行いつつ、昨今の AI 処理やデータ処理で用いられる疎行列計算を高速化するための CGRA（Coarse-Grained Reconfigurable Array、粗粒度再構成可能アーキテクチャ）構成やソフトウェアツールチェーンなどの研究に従事。昨今の AI で主流となっている LLM（Large Language Model、大規模言語モデル）アクセラレータシステムの研究を実施した。

同時に、音声認識を対象とした AI プロセッサ関連技術の研究も行い、2024 年 4 月に開催したワークショップでの議論をもとに研究を推進した。また、グラフニューラルネットワーク（GNN）の活用による ASIC（Application Specific Integrated Circuit、特定用途向け集積回路）設計の容易化に関して、ASIC 設計に関連する回路や RTL といった表現は自然にグラフ構造をとるため、グラフ構造のデータを処理する GNN の活用が多数提案されていることから、ASIC 設計データの特徴を抽出しやすい GNN を用いて、論理合成や配置配線の結果予測を短時間で行い、設計期間短縮を実現するためのフレームワークを提案した。さらに、その実現に向けた学習データセットの作成を進めた。

<得られた成果>

新たなシリアル FFT（Fast Fourier Transformation）を用いた AI プロセッサ方式を考案し、その一部成果について国際学会 2024IEEE APC-CAS にて口頭発表を行った。

② 国際頭脳循環の促進にかかる成果

<実施したこと>

2024 年 4 月にキックオフシンポジウムを米国スタンフォード大学他で開催し、3 日間の日程で設計高効率化並びに AI プロセッサ設計に関して 10 件超の口頭発表を行った。また、2024 年 9 月から 3 名の若手研究者がスタンフォード大学に滞在し研究に従事している。

<得られた成果>

新たな共同研究先を開拓できた。

2. 研究実施体制

研究テーマ	中心となる研究者氏名	所属機関・部署・役職名
研究テーマ1	小菅 敦丈	東京大学大学院・工学系研究科附属システムデザイン研究センター・講師
研究テーマ2	小島 拓也	東京大学大学院情報理工学系研究科・システム情報学専攻・助教
研究テーマ1/2	Priyanka Raina	Assistant Professor, School of Engineering, Stanford University
研究テーマ3	門本 淳一郎	東京大学大学院・情報理工学系研究科電子情報学専攻・講師

3. 代表的な業績（原著論文、プレスリリース、表彰など）

•A. Kawada, K. Kobayashi, J. Shin, R. Sumikawa, M. Hamada, A. Kosuge, "A 250.3uW Versatile Sound Feature Extractor Using 1024-Point FFT 64-ch LogMel Filter in 40nm CMOS," in IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Nov. 2024.