「イノベーション創発に資する人工知能基盤技術の創出と統合化」
平成 28 年度採択研究代表者

H29 年度
実績報告書

篠田 浩一

東京工業大学情報理工学院
教授

社会インフラ映像処理のための高速・省資源深層学習アルゴリズム基盤

§1. 研究実施体制

(1)篠田グループ
① 研究代表者:篠田 浩一 (東京工業大学情報理工学院 教授)
② 研究項目
・知識の構造を活用した高速な深層学習アルゴリズム

(2)松岡グループ
① 主たる共同研究者:松岡 聡(東京工業大学情報理工学院 教授)
② 研究項目
・ノード間の通信処理を削減するための高並列アルゴリズムと資源スケジューリングによる全体最適化

(3)村田グループ
① 主たる共同研究者:村田 剛志(東京工業大学情報理工学院 准教授)
② 研究項目
・リアルタイム認識・解析のための Deep Net 構造のコンパクト化アルゴリズム

(4)横田グループ
① 主たる共同研究者:横田 理央(東京工業大学学術国際情報センター 准教授)
② 研究項目
・個々の計算ノードにおける計算量を削減するための行列構造化アルゴリズム

§2. 研究実施の概要

この研究期間内では、昨年度構築した評価プラットフォームを活用して、各グループが各々の課題を遂行し、年度内における目標をほぼ達成した。全体では、Small Phase の目標に対し、だいたい半分の進捗である。また、様々な計算機環境やツールの性能比較・検討を行い、この課題全体での統一化を進めた。さらに、様々なハードウェア環境での機械学習の実行をシミュレートできるプロトタイピングフレームワークの開発に着手した。アーキテクチャー・アルゴリズムの優劣を実装前に判断できることを目的とする。以下、各グループ個別の進捗を述べる。

横田グループでは、2016 年度に低ランク近似を用いた行列構造化アルゴリズムによる深層学習の演算量低減を行い、メモリ消費量、演算量を 2 分の 1 に低減した[1]。2017 年度には、この低ランク近似の技術を深層学習の最適化手法である K-FAC 法と組み合わせた。K-FAC 法はクロネッカー因子分解によるフィッシャー情報行列の近似を行うことで自然勾配法の高速化を実現するものである。従来の確率的勾配降下法と比べて収束が 10 倍速く、また、学習効率がバッチサイズの増大の影響を受けない。

松岡グループでは、まず、スパコンなどの高並列環境における巨大行列に対する畳み込み演算の実装において、行列を分割して各部分に対し別々のアルゴリズムを適用することで、従来の単一アルゴリズムを用いる方法よりも、1.2〜1.6 倍の高速化を実現した[2]。また、通信量の削減について、昨年度までの通信時の浮動小数点精度を固定する方法に代えて、通信時の浮動小数点数の精度を動的に選択する手法を開発し、認識性能を劣化させることなく 2.3〜2.5 倍の高速化を達成した。

篠田グループでは、まず、「構造的教師ラベルを用いた学習」においては、映像からのイベント検出において、イベントを構成するコンセプト間の関係を利用することで、学習データが少なくても安定して深層学習を行う手法を構築した。NIST TRECVID ワークショップのマルチメディアイベント検出タスクで世界 2 位の性能を達成した。「因果関係を抽出して用いる能動学習」においては、時系列データに対する深層学習法である CTC 法と統計言語モデルを組み合わせて因果関係をモデル化する手法を提案した。国際会議 ACM マルチメディアでその成果を発表した[3]。

村田グループでは、「Deep Net の局所的なサイズ圧縮による 2/3 程度のサイズ圧縮」と「Deep Net の大域的な構造の最適化による 1/4 程度のサイズ圧縮」を目標に研究を行った。枝刈り、量子化、符号化を組み合わせて、従来手法を改良することにより、認識精度を悪化させずにサイズを 1/90 にすることができ、目標を大幅に超える性能を達成した。一方、Deep Net の大域的な構造の最適化は良い結果が得られず、Deep Net などの密なモデルには有効な手法ではないことがわかった。

[1] K. Osawa, R. Yokota, “Evaluating the Compression Efficiency of the Filters in Convolutional Neural Networks”, In Proceedings of the 26th International Conference on Artificial Neural Networks, pp, 459–466 (2017).
[2] Y. Oyama, T. Ben-Nun, T. Hoefler, S. Matsuoka, “μ-cuDNN: Accelerating Deep Learning Frameworks with Micro-Batching”, https://arxiv.org/abs/1804.04806.

[3] M. Lin, N. Inoue, and K. Shinoda, “CTC Network with Statistical Language Modeling for Action Sequence Recognition in Videos”, In Proceedings of the on Thematic Workshops of ACM Multimedia 2017, pp.393-401, 2017