

令和 5 年 6 月 13 日現在

機関番号：12608

研究種目：基盤研究(B)（一般）

研究期間：2020～2022

課題番号：20H04165

研究課題名（和文）異種アーキテクチャ並列環境におけるスケーラブルな機械学習基盤ソフトウェア技術

研究課題名（英文）Scalable System Software for Machine Learning on Heterogeneous Parallel Computing Environments

研究代表者

遠藤 敏夫（Endo, Toshio）

東京工業大学・学術国際情報センター・教授

研究者番号：80396788

交付決定額（研究期間全体）：（直接経費） 15,870,000円

研究成果の概要（和文）：深層学習は多くの計算量を必要とするためにGPUなどの高性能プロセッサを多数必要とする。しかしその高性能を引き出すためには複雑な計算機アーキテクチャの特性を考慮する必要があり、新規学習アルゴリズムの取り入れに障害となる。本研究では高性能とソフトウェア開発コスト低減の両立を対象とした。統合成果の一つとして、多数GPUを持つスパコン上の学習を改善するためのハイブリッド並列学習フレームワークを構築した。

研究成果の学術的意義や社会的意義

LLMの隆盛に代表されるように、機械学習処理には社会を変革する力があると考えられるが、アルゴリズムの改良と大規模な計算機アーキテクチャを活用するための技術の双方が必要である。本研究では主に後者の立場から、GPUなどのプロセッサ内の処理効率化と、多数GPU搭載計算機をよどみなく活用するフレームワーク等を実現した。これら基盤技術により、LLMのさらなる改善への応用が期待される。

研究成果の概要（英文）：Deep learning is computationally intensive and requires lots of high performance processors such as GPUs. However, in order to achieve high performance, it is necessary to consider the characteristics of complex computer architectures, which hinders the incorporation of new learning algorithms. This research targets both high performance and software development cost reduction. As one of the integrated results, we constructed a hybrid parallel learning framework to improve learning on a supercomputer with many GPUs.

研究分野：高性能計算ソフトウェア

キーワード：深層学習 GPU メモリ階層 DSL

## 1. 研究開始当初の背景

自動運転や自動翻訳などの社会構造を変えうる人工知能技術のキーテクノロジーとして、深層学習を中心とする機械学習の研究が急速に進展している。機械学習・深層学習への参入障壁も下がっており、それは pyTorch, TensorFlow などの深層学習フレームワークなどの進展によるところが大きい。そしてそれらの多くは Python 言語・NumPy/Pandas ライブラリを中心とする迅速なプロトタイピングを可能とする基盤ソフトウェア技術により支えられており、日々新たな学習手法が提案され、一部は即時にオープンソフトウェア化され、別手法の基盤になるといった、エコシステムを形成している。

機械学習において大規模データを操作する必要がある場合、演算速度や操作可能なデータ規模のスケール向上のために高速プロセッサ群や大容量メモリの活用が必要となり、そのための研究開発も広く行われている。たとえば、深層学習のブームのきっかけの一つは画像認識コンテスト ILSVRC2012 におけるトロント大学チームの優勝であるが、このチームでは深層学習のために多数の GPU を持つ計算機群を用いて数週間の学習演算を行った。CPU や GPU に加え機械学習・深層学習に特化したプロセッサ(Google TPU, Cerebras WSE など)の提案が行われているなど、プロセッサアーキテクチャは異種・多様化に向かっている。またメモリ分野においても、キャッシュ・メインメモリ・ハードディスクと言った古典的記憶階層の間を埋める技術として、High Bandwidth Memory (HBM), Intel 3D-XPoint などの新規デバイスがすでに利用可能となっている。

上記のハードウェア動向における、異種プロセッサの高性能を機械学習のエンドユーザから活用するために様々な基盤ソフトウェア・ライブラリが開発されている。特に NVIDIA GPU のための cuDNN, cuBLAS や Intel CPU のための MKL-DNN、ARM Performance libraries など、プロセッサベンダーによる高度に最適化されたものが先行するケースが多い。

これらのライブラリ等により、各プロセッサの特性を最大限に引き出すことができる一方で、日々提案される新しい学習手法やアルゴリズムが常にそれらを活用できるとは限らない。特に機械学習分野においては新しい手法の利害得失を考察して改良を繰り返すフェーズが長く、プロトタイプを低コストで実装する必要がある。

もう一つの軸として、単独プロセッサ/ノードで賄いきれない大規模学習のためには分散並列技術が必要となる。そのためのソフトウェア環境も複数存在し、TensorFlow や Chainer のフレームワークの拡張、Dask などの NumPy/Pandas をベースにした Python ライブラリなど、アプローチも複数存在する。一方で、分散並列技術と上記に述べたような異種プロセッサや新規メモリデバイスの同時活用は遅れる傾向にある。

以上の議論から、本研究の問題意識は、爆発的多様化するプロセッサ・メモリアーキテクチャ上で、性能・データ規模ともにスケラブルな機械学習ソフトウェアの開発コストを抑制し、その継続的な進化を可能とすることにある。

## 2. 研究の目的

本研究の目的は、莫大な演算量やメモリを必要とする機械学習を、プロセッサ・メモリ階層とも複雑化・異種混合化し、今後も変貌し続ける計算機アーキテクチャ上で高性能・大規模に実行可能とするソフトウェア技術の確立である。機械学習・深層学習分野においては新たな手法・アルゴリズムの進展が日常的という前提のもと、proof of concept 実装と検証のループを含む開発コスト軽減に大きく寄与することをねらう。

## 3. 研究の方法

これまでに述べた研究目的に向けて、主に下記の項目の研究を行う計画とした。

- (a) 演算カーネルレベルのマルチコア/SIMD 並列性の活用:爆発的多様化するアーキテクチャ上でソフトウェア開発コストを低減するうえで、適切なプログラミングモデルの探求が本項目に含まれる。理想的には、逐次言語に近い形でアルゴリズムを記述できることが望ましい。この項目の文脈で注目したアプローチの一つは Halide に代表される配列・行列操作記述のためのドメイン特化型言語(DSL)である。このような DSL においては記述可能な演算がより限定されており、これは長所でも短所でもある。並列化可能な部分の認識がより容易であり最適化を行いやすい一方で、記述できるカーネルの種類が減少するのは短所である。このような制限とカーネル記述の自由性を見極めおよび速度性能への影響などを評価する。
- (b) 水平・垂直方向の演算・データスケジューリング:項目(a)は主にカーネルレベルを対象とするが、より大規模な範囲のソフトウェアの演算・データおよびその移動を、複雑な計算機アーキテクチャにマッピングするソフトウェア技術を研究する。
- (c) 大規模機械学習ソフトウェアとしての技術統合:項目(a) (b)の成果を統合し、機械学

習ソフトウェアとして実現する。到達目標としては、深層学習フレームワーク(PyTorch, Tensorflow など)上を用いた学習アルゴリズムや、行列分解を中心とする推薦アルゴリズムで、主アルゴリズムは Python で記述されたものに対して、異種プロセッサ・メモリ・複数ノードをスケラブルに活用可能な形で実行を行う。

#### 4. 研究成果

##### 研究項目(a) 演算カーネルレベルのマルチコア/SIMD 並列性の活用：

まずドメイン特化型言語(DSL)Halideによる開発コストの低減とCPU/GPU上の高性能のトレードオフに関する研究を行った。研究対象のカーネルとして、畳み込み演算の基本ともなるステンシルカーネルについて、キャッシュ利用効率を向上させるために反復をまたぐブロッキング技術(時間ブロッキング)を適用したものを挙げた。これは依存関係を満たすために手作業の実装ではループ構造や範囲指定が複雑になるカーネルである。Halideを用いたプロトタイプ実装と手作業実装をCPU/GPU上で性能比較することにより、低開発コストと高性能が両立可能であることを示した。いっぽうHalide利用の障害の一つとしてコンパイルの複雑さなどが挙げられる。他の並列プログラミング環境として、C++ラムダ式に基づいたNVHPCコンパイラやKokkosライブラリによるアルゴリズム記述にも注目した。これによりユーザによる各点の演算式の記述とループ構造の記述の分離が可能である。ステンシルカーネルへの時間ブロッキング適用を対象に初期実装・評価を行った。

また、映像検出タスクの一種であるSingle Shot Detector(SSD)を対象とし、実行速度へ最も影響の大きいカーネルとしてソートとNMS処理を洗い出した。当時のcuDNNなどの既存ライブラリのみでは十分な性能を得るのに不足であると判断したため、必要なカーネルをCUDAで実装(現在は手作業)し、先行研究による実装よりも高速な実装を実現した。

さらに、演算カーネルレベルにおけるGPU性能を引き出すために、GPU Tensor coreを用いたステンシル演算への時間ブロッキングへの統合および、GPUカーネル融合技術などを用いたAttentionカーネルの研究を行い、開発コスト低減技術との統合を見込んでいる。

##### 研究項目(b) 水平・垂直方向の演算・データスケジューリング：

マルチGPU環境における高性能深層学習に向けて、データ並列とパイプライン並列を併用したハイブリッド並列のさらなる改良手法を提案した。負荷バランスとAll-reduce対象プロセス数の低減が見込める一方、プロセス間の通信パターンが複雑になるために通信処理の順序やスレッド数が性能に大きく影響することを示した。その一環としてPyTorch内部の通信処理部分の改良を行うことにより、複数種類の通信が発生するハイブリッド並列においてアイドル時間を抑制した。

並行して、ニューラルネットワーク構造およびハードウェア構造を考慮したマッピングの改良の研究を進めている。並行して、音声認識手法であるAugment Adversarial Training(AAT)を主対象として、マルチGPUにおけるハイブリッド並列アルゴリズムを検討し、プロトタイプ実装を行った。

またこの項目に関して、比較的複雑なカーネル間依存関係を持つ大行列のコレスキー分解処理を対象とし、開発コスト低減と性能の両立について研究した。複数ノード上でPython Daskスケジューリングライブラリの活用などにより通信対象を明記せずにスケラビリティを得た。

以上の研究内容は演算コアおよびCPU/GPU上のメモリを用いるものであったが、ストレージデータのスケジューリングに関する研究を以下のように行った。GPU上の学習処理に基づいた高性能アプリケーションとして、AlphaFoldなどのタンパク質構造推測ソフトウェアが注目されている。このようなソフトウェアでは、GPU/CPU上の高性能並列処理に加え、2.2TBの大規模データへのアクセスを必要とする。そのデータアクセスがAlphaFoldの実行時間に多大な影響を及ぼすことを示し、TSUBAME上のハードディスクベースの共有ストレージ・ローカルSSD・RAMディスクといった性質の異なる異種ストレージに適切に配置する手法を提案・高速化を実証した。

##### 研究項目(c) 大規模機械学習ソフトウェアとしての技術統合：

この項目においては、項目(b)に述べた改良ハイブリッド並列手法を基に、大規模深層学習が可能なフレームワークをPyTorch上に実装した。このフレームワーク上で、Visual Transformer(層数360のViT-H-14)を含むモデルの学習処理の評価実験をTSUBAMEスパコンの多数GPU上で行い、単純なデータ並列に比べて高速化することを実証した。一方でパイプライン並列の導入に伴う学習精度の低下が観測された。これを改善するために、学習途中でデータ並列とハイブリッド配列を切り替える手法(Switch parallel方式)を提案・実装し、速度と学習精度の両立を実現した。

以上の研究の一部について、産総研・東工大 実社会ビッグデータ活用 オープンイノベーションラボラトリおよび産総研デジタルアーキテクチャ研究センターとの共同で実施した。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 Wang Chenyu, Endo Toshio, Hirofuchi Takahiro, Ikegami Tsutomu	4. 巻 5
2. 論文標題 Pyramid Swin Transformer: Different-Size Windows Swin Transformer for Image Classification and Object Detection	5. 発行年 2023年
3. 雑誌名 Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications	6. 最初と最後の頁 583-590
掲載論文のDOI（デジタルオブジェクト識別子） 10.5220/0011675800003417	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Aikawa Hiroki, Endo Toshio, Yuki Tomoya, Hirofuchi Takahiro, Ikegami Tsutomu	4. 巻 -
2. 論文標題 Efficient Stencil Computation with Temporal Blocking by Halide DSL	5. 発行年 2022年
3. 雑誌名 proceedings of 20th IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA)	6. 最初と最後の頁 870-877
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ISPA-BDCIcloud-SocialCom-SustainCom57177.2022.00116	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Wang Chenyu, Endo Toshio, Hirofuchi Takahiro, Ikegami Tsutomu	4. 巻 -
2. 論文標題 Speed-Up Single Shot Detector on GPU with CUDA	5. 発行年 2022年
3. 雑誌名 proceedings of 23rd ACIS International Summer Virtual Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD2022-Summer)	6. 最初と最後の頁 89 ~ 106
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-3-031-19604-1_7	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計6件（うち招待講演 0件／うち国際学会 0件）

1. 発表者名 相川 洋貴, 遠藤 敏夫, 幸 朋矢, 広瀬 崇宏
2. 発表標題 時間ブロッキングを用いたステンシル計算のHalide言語による高性能実装と評価
3. 学会等名 並列/分散/協調処理に関するサマワークショップ(SWoPP2021)
4. 発表年 2021年

1. 発表者名 細木 隆豊, 野村 哲弘, 遠藤 敏夫
2. 発表標題 GPUクラスタにおけるハイブリッド並列DNN学習のボトルネック分析と改良
3. 学会等名 並列/分散/協調処理に関するサマールークシヨツプ(SWoPP2021)
4. 発表年 2021年

1. 発表者名 瓜生 侑, 遠藤 敏夫
2. 発表標題 ラムダ式を用いる移植性の高い並列プログラムの実装とCPU・GPU上の評価
3. 学会等名 並列/分散/協調処理に関するサマールークシヨツプ(SWoPP2022)
4. 発表年 2022年

1. 発表者名 大沢 泰生, 遠藤 敏夫, 野村 哲弘
2. 発表標題 タンパク質構造解析システムAlphaFoldの実行時ファイルステージングを用いた高速化
3. 学会等名 並列/分散/協調処理に関するサマールークシヨツプ(SWoPP2022)
4. 発表年 2022年

1. 発表者名 細木 隆豊, 遠藤 敏夫, 広淵 崇宏, 池上 努
2. 発表標題 負荷分散を改善したハイブリッドパイプライン並列深層学習手法
3. 学会等名 並列/分散/協調処理に関するサマールークシヨツプ(SWoPP2022)
4. 発表年 2022年

1. 発表者名 丸山 翼, 池上 努, 遠藤 敏夫, 広淵 崇宏
2. 発表標題 機械学習を用いた音声処理に向けたデータ拡張手法の研究
3. 学会等名 音響・超音波サブソサイエティ合同研究会
4. 発表年 2022年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

当研究課題の成果発表「時間ブロッキングを用いたステンシル計算のHalide言語による高性能実装と評価」により、相川 洋貴（当時修士課程学生）が情報処理学会HPC研究会学生優秀発表賞を受賞した。

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関