

戦略的創造研究推進事業 CREST
研究領域「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」
研究課題「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」

研究終了報告書

研究期間 2019年4月～2022年3月
(追加支援により、2023年3月まで延長)

研究代表者: 篠田浩一
(東京工業大学情報理工学院、教授)

§ 1 研究実施の概要

(1) 実施概要

機械学習と高性能計算の研究者が密に連携した Co-Design の枠組みを活用して現在より 10 万倍速い超高速な深層学習基盤を構築し、その上で大量の高精細映像の実時間解析を実現して安心・安全な社会作りに貢献することを目的としている。

機械学習においては、二次最適化、自己教師付き学習、ニューラルモデル構造探索、グラフニューラルネットワークなどの研究を行った。クロネッカー因子分解を用いた二次最適化の高速近似解法はこれまでその有効性が経験的にしか示されてこなかったが、層の数が無限大の極限では厳密な自然勾配法と同じ収束率をもつこと、深層学習のような超高次元モデルでもノイズを加えながら勾配法を実行することで大域的最適解に収束することを理論的に証明した。さらに、グラフニューラルネットワークが多層になると性能が劣化する問題の理論的解明を行い、その問題を回避したマルチスケール GCN 法を提案した。

また、スーパーコンピュータ「富岳」における 10 万ノード規模の学習アルゴリズムの超並列化に向けて、3D CNN におけるモデル並列化を推進するとともに、言語モデルや脳シミュレーションのアプリについても「富岳」の評価環境を用いた並列処理性能の予備評価を進めた。さらに、Transformer の超並列分散学習アルゴリズムを開発し、「富岳」にて 80000 ノードでの並列学習を実現している。一方で、物体検出や領域分割などの応用に向けて FPGA 実装を行い、物体検出応用に関して CPU や GPU よりも高速な処理を実現した。学習後のモデルを FPGA に効率よくポータリングするための Co-Design 環境を開発した。

新型コロナウイルスの感染リスクの評価を目的として 2021 年 1 月 4 日に国立競技場で開催された Jリーグヴァンカップ決勝戦の観戦映像データを取得した。また、これまでに作成してきた学習モデルを用いて取得したデータを処理することで、顔が識別できない程度の解像度の画像でも人の領域を抽出できることが分かった。

次いで、研究ビジョンとその達成状況について述べる。当該課題の研究ビジョンのステートメントは以下の通りである。

「現在より 10 万倍速い超高速な深層学習を実現し、AI の研究開発にパラダイムシフトをもたらす。特に大量の高精細映像の実時間解析を実現し、安心・安全な社会作りに貢献する。」

達成状況は、「富岳」を用いた 8 万倍速い超高速な深層学習、リアルタイムの 100 倍のビデオ画像処理等、ステートメントの目標を十分に達成するソーシャルインパクトがある結果を得られた。

なお、研究ビジョンの達成指標は以下の通りである。

1. 映像規模のモデルに対し 10 万ノード程度の並列学習を実現
2. モデルを小型化し FPGA で実時間×3 倍の速度で認識
3. 学習から認識までのモデルシミュレーションツールを公開

これに対し、2022 年 3 月までに以下を達成した。

1. 「富岳」にて 8 万ノードでの並列学習を実現
2. 通常のビデオ画像でリアルタイムの 100 倍
3. 学習・認識各々のシミュレーションツールを開発

※超並列化の更なるスケールアップと映像応用への展開を目標として、1 年間研究期間を延長し、1) スーパーコンピュータ「富岳」を用いた 10 万ノード規模の Vision Transformer 超並列学習、2) 映像を用いた質問応答システム、3) 少量データによる人間の動作生成の研究を実施した。

1) では超並列分散処理で課題となる、メモリ資源の最適配置方法、学習時における異常値発生メカニズムの解析、性能を保ちつつデータのランダムアクセスを制限する手法、「富岳」の

CPU 演算の高速化アルゴリズム、圧縮ファイルからデータを直接読み込む手法、確率勾配降下法における陰的正則化モデリングなどの開発を行い、82944 ノード数での超並列学習で単体計算ノードに比べ約 8 万倍の高速化を実現した。2) では物体検出器を質問応答の内容に適応させることで Vision Transformer の質問応答性能を向上させる手法を開発し、画像分野でトップクラスの国際会議 WACV2023 に採択された。3) では、深層暗黙表現を用いて人間の動作を生成する Transformer を開発し、動作の正確性と多様性の評価で世界最高性能(発表当時)を達成した。この成果も画像分野のトップ会議 ECCV2022 に採択された。

(2) 顕著な成果

<優れた基礎研究としての成果>

1. クロネッカー因子分解による二次最適化が自然勾配法と同等の収束性をもつことを証明

概要:

これまでその有効性が経験的にしか示されてこなかったクロネッカー因子分解を用いた 2 次最適化の高速近似解法が、層の幅が無限大の極限では厳密な自然勾配法と同じ収束率をもつことを理論的に証明した。その結果は採択率 1.1%の NeurIPS2020 の口頭発表論文として採択された。

2. 深層学習が過学習しない理由の解析

概要:

深層学習がなぜサンプルサイズより大きなパラメータ数を持ちつつも過学習しないかをニューラルネットワークの圧縮可能性の視点から解析した。既存の研究では圧縮したネットワークの汎化誤差を解析していたが、本研究では潜在的に圧縮可能なネットワークが汎化できることを示し、既存の評価よりも厳密な評価を得た。

3. マルチモーダル感情認識

概要:

深層学習によるマルチメディア認識では複数のモードの情報が複雑に絡み合っており、それらを解きほぐす(disentanglement)する手法の開発が重要である。ここでは、音声为例にとり、そこから、音韻(言語)や話者性に関する情報を分離することで、それらの違いに頑健な感情認識手法を開発した。従来の感情認識手法を上回る性能を得た。この手法は、映像における人間の動作の認識などにも容易に適用が可能である。

<科学技術イノベーションに大きく寄与する成果>

1. クロネッカー因子分解による二次最適化を用いた超並列分散学習

概要:

ディープニューラルネットワークの大規模な分散学習では、バッチサイズの増加により汎化性能が劣化する現象が知られている。これに対し、スケーラブルで実用的な自然勾配降下法を提案した。クロネッカー因子分解やブロック対角化近似などの手法を用いることで、一次最適化法で学習したモデルと同様の汎化性能を達成するのにかかる時間を加速した。

2. FPA による画像認識の高速化

概要:

分担者中原が起業した、東工大発スタートアップにおいて、CREST の研究成果の一部を活用して FPGA などの組み込みデバイスに AI を実装・開発するサービスを提供する事業を手掛けている。一例として、マルハニチロと提携した養殖場の魚を自動カウントするシステムを実装し、販売している。その成果の一部は日本経済新聞に掲載されている。

3. 新型コロナ対応のための人の動き解析

概要:

Jリーグや NPB と連携して高解像度カメラを用いてスタジアムの客席を個人が特定できない程度の大きさで撮影し、深層学習を用いて人物領域を抽出し、抽出した領域のマスクの着用、顎マスク、未着用のクラス分類を行う手法を実装し、実際のスタジアムで実運用し評価した。人物領域の抽出は現在スタジアム内にいる人数が把握できることから分散退場の効果を検証することが可能になった。

<代表的な論文>

1. Kazuki Osawa, Yohei Tsuji, Yuichiro Ueno, Akira Naruse, Chuan-Sheng Foo, Rio Yokota, Scalable and Practical Natural Gradient for Large-Scale Deep Learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 2020; PP:10.1109/TPAMI.2020.3004354.

概要:

これまでその有効性が経験的にしか示されてこなかったクロネッカー因子分解を用いた2次最適化の高速近似解法が、層の幅が無限大の極限では厳密な自然勾配法と同じ収束率をもつことを理論的に証明した。その結果は採択率 1.1%の NeurIPS2020 の口頭発表論文として採択された。

2. Taiji Suzuki, Hiroshi Abe, Tomoaki Nishimura: Compression based bound for non-compressed network: unified generalization error analysis of large compressible deep neural network, ICLR2020, spotlight.

概要:

深層学習がなぜサンプルサイズより大きなパラメータ数を持ちつつも過学習しないかをニューラルネットワークの圧縮可能性の視点から解析した。既存の研究では圧縮したネットワークの汎化誤差を解析していたが、本研究では潜在的に圧縮可能なネットワークが汎化できることを示し、既存の評価よりも厳密な評価を得た。

3. Mariana Rodrigues Makiuchi, Kuniaki Uto, Koichi Shinoda, Multimodal emotion recognition with high-level speech and text features, Proceedings of Automatic Speech Recognition and Understanding (ASRU) 2021, 10.1109/ASRU51503.2021.9688036

概要:

深層学習によるマルチメディア認識では複数のモードの情報が複雑に絡み合っており、それらを解きほぐす(disentanglement)する手法の開発が重要である。ここでは、音声为例にとり、そこから、音韻(言語)や話者性に関する情報を分離することで、それらの違いに頑健な感情認識手法を開発した。従来の感情認識手法を上回る性能を得た。この手法は、映像における人間の動作の認識などにも容易に適用が可能である。

4. Pablo Cervantes, Yusuke Sekikawa, Ikuro Sato, Koichi Shinoda: Implicit Neural Representations for Variable Length Human Motion Generation, Proceedings of 17th European Conference of Computer Vision (ECCV2022), pp. 356-372, https://doi.org/10.1007/978-3-031-19790-1_22..

※ 1年追加支援時の成果

概要:

異常検知のための人間の動作の認識では、稀な動作を高い精度で認識することが重要である。そのために動作モデルの学習に用いるデータを生成するデータ増強が有望である。ここでは、深層暗黙表現(Deep Implicit Representation)を用いた動作生成を提案した。固定長の数値ベクトルから、様々な継続時間長をもつ動作を生成することが可能である。動作の多様性と動作の判別性の両方の基準で世界最高の性能(発表時点)をもつ。

§2 研究実施体制

(1) 研究チームの体制について

① 「篠田」グループ

研究代表者：篠田 浩一 (東京工業大学情報理工学院 教授)

研究項目

- ・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

② 「松岡」グループ

主たる共同研究者：松岡 聡 (理化学研究所計算科学研究センター センター長)

研究項目

- ・10万ノード超並列処理と資源スケジューリング

③ 「大西」グループ

主たる共同研究者：大西 正輝 (産業技術総合研究所人工知能研究センター 研究チーム長)

研究項目

- ・実社会応用における評価

④ 「横田」グループ

主たる共同研究者：横田 理央 (東京工業大学情報理工学院 准教授)

研究項目

- ・確率的勾配降下法に代わる二次最適化

⑤ 「村田」グループ

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)

研究項目

- ・グラフ深層ニューラルネットによる知識処理との融合

⑥ 「中原」グループ

主たる共同研究者：中原 啓貴 (東京工業大学工学院 准教授)

研究項目

- ・FPGA 実装向け深層学習回路

⑦ 「鈴木」グループ

主たる共同研究者：鈴木 大慈 (東京大学大学院情報理工学系研究科 准教授)

研究項目

- ・深層学習理論(汎化誤差、圧縮率、収束率)

※1年追加支援時の体制

① 「篠田」グループ

研究代表者：篠田 浩一（東京工業大学情報理工学院 教授）

研究項目

- ・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

② 「松岡」グループ

主たる共同研究者：松岡 聡（理化学研究所計算科学研究センター センター長）

研究項目

- ・10万ノード超並列処理と資源スケジューリング

③ 「横田」グループ

主たる共同研究者：横田 理央（東京工業大学情報理工学院 准教授）

研究項目

- ・確率的勾配降下法に代わる二次最適化

(2) 国内外の研究者や産業界等との連携によるネットワーク形成の状況について

1. NEDO の「次世代人工知能・ロボットの中核となるインテグレート技術開発/人工知能技術の適用領域を広げる研究開発/人工知能技術の導入加速化技術/自動機械学習による人工知能技術の導入加速に関する研究開発」(代表 大西正輝 2020年度～2024年度)を開始した。本 CREST で培った深層学習技術をハイパーパラメータ最適化やニューラルアーキテクチャサーチなどの AutoML の技術で導入を加速化させる研究開発である。
2. AMED の未来医療を実現する医療機器・システム研究開発事業 ICT を活用した診療支援技術開発プロジェクトに加わり「表情・音声・日常生活活動の定量化から精神症状の客観的評価をリアルタイムで届けるデバイスの開発」の委託研究を行った(～2018年度)。本 CREST において開発した少量データからの深層学習の技術を用いて、音声を入力とする認知症診断を実現した。
3. 科研費基盤研究 A「大規模脳波データとキャリブレーションレスモデルの構築による意思伝達 BMI の実現」(代表 田中聡久教授・農工大, 2020年度～2023年度)に分担者として参加している。本 CREST において開発した少量データからの深層学習の技術を用いて、脳皮質電位を用い想起音声(音声を出すことを思い描くだけで実際に音声は出さない)を認識する方式の研究である。
4. 農水省の農林水産研究推進事業委託プロジェクト研究「現場ニーズ対応型プロジェクト」 「品種多様性拡大に向けた種子生産の効率化技術の開発」(2020年度～2022年度)を分担している。農研機構が中心となっているプロジェクトであり、本 CREST で培った映像技術を用いて、ドローンで撮像した水田・畑などの映像を用いて、穀物の生育状態を観測・解析し、生産効率向上につなげる。
5. 理化学研究所計算科学研究センターに研究チーム「高性能人工知能システム研究チーム」(松岡 聡チームリーダー)が本プロジェクトを基に設立された。
6. 株式会社富士通研究所と東京工業大学の松岡チームのメンバーで共同研究「次世代コンピュータシステムのソフト・ハードアーキテクチャと適用アプリに関する研究」(2020.04-2021.03)を実施している。当該研究では次世代アーキテクチャにおける各種アプリの適用可能性を検証しており、本 CREST で対象としている人工知能領域のアプリケーションは中でも核をなすアプリケーションとして重視し、積極的に研究を行っている。
7. 株式会社富士通(事業部)、株式会社富士通研究所、理化学研究所計算科学研究センター、東京工業大学で4者 MOU を締結し、スーパーコンピュータ「富岳」のアーキテクチャである A64fx における機械学習アプリケーションの性能検証・高速化を行う「DL4Fugaku プロジェクト」を実施している。

8. 産業技術総合研究所の研究開発計画「Easily-Buildable AI」 with ABCI 2.0(2020 年度-2024 年度)において、機械学習アプリケーションのデータおよびモデルの効率的な管理プラットフォームについての研究を本 CREST の成果をベースとして行っている。
9. 東京工業大学情報理工学院は 2020 年 4 月に共同研究講座「DENSO IT LAB 認識・学習アルゴリズム共同研究講座」を設立した(代表:篠田浩一)。デンソーアイティラボラトリーは従来から本 CREST の研究協力者であり、ともに研究を行ってきた。今後、本 CREST の研究者も含めた 7 名の教員と 2 名の特任教員のチームで、次世代モビリティに向けた共同研究開発を行う。そこでは、本 CREST で培った映像処理向けの深層学習技術が基盤となる。
10. 研究参加者である二反田篤史(東京大学大学院情報理工学系研究科特任助教)の研究提案「深層学習の潜在的正則構造の理解に基づく学習法の安定化と高速化」が JST さきがけ「数学と情報科学で解き明かす多様な対象の数理構造と活用」(研究総括:坂上貴之)に採択された。内容は、深層学習の理論研究を主に最適化の観点から推進するものである。深層学習の効率的計算機実装およびその理論研究を推進する本 CREST の派生的課題と言え、両課題の相互作用によって研究の活性化が期待できる。
11. JST CREST 研究領域「数学・数理科学と情報科学の連携・融合による情報活用基盤の創出と社会課題解決に向けた展開」(研究総括:上田修功)における 2020 年度採択研究課題「数理知能表現による深層構造学習モデルの革新」(代表:福水健次)に鈴木が主たる共同研究者として参画することになった。これは本 CREST に比べて大きく理論寄りの研究課題であるが、本 CREST における実践を主とする研究と知識を相互にフィードバックさせることで、両課題の研究が加速されることが期待される。
12. 2021 年度より、JST 共創の場形成支援プログラム(COI-NEXT)本格型「誰もが参加し繋がることでウェルビーイングを実現する都市型ヘルスコモンズ共創拠点」に篠田が研究担当者として参加することになった。本 CREST で得られた深層学習の成果、特に、身振り認識・感情認識などの手法を認知症・うつ病などの精神疾患の早期発見のために活用する。