

報道の解禁日(日本時間)

(テレビ,ラジオ,インターネット):2024年2月20日(火)18時00分

(新聞):2024年2月21日(水)付朝刊

2024年2月15日

記者會、記者クラブ 各位

低価格コンピュータで学習&推論可能な AI 基盤モデルを開発 軽量かつ既存の大規模 AI モデルと同程度の性能を達成

【本研究のポイント】

- ・低価格なコンピュータ(GPU)で物体認識などの推論が可能なマルチモーダル AI 基盤モデルを開発
- ・効率の良いマルチモーダル AI の学習方法の提案
- ・軽量でありながら、既存の大規模 AI モデルと同程度の性能を確認

【研究概要】

岐阜大学工学部の加藤邦人教授と日本車輛製造株式会社の共同研究チームは、物体認識、物体検知など AI の適用先の範囲を適度に限定することで、軽量ながらも既存の大規模モデルに匹敵する性能を持つ汎用マルチモーダル AI の基盤モデルを開発しました。この AI モデルでは、学習と推論を低価格なコンピュータ(GPU)上で効率的に行うことが可能です。また、この研究は AI の学習コストを大幅に削減するという重要な進歩を示しました。

この成果は、2024年2月20日にコンピュータービジョンの国際ワークショップ IW-FCV で発表される予定です。

【研究背景・成果】

言語や画像など、多様な情報を理解するマルチモーダルモデル¹⁾は、画像認識や画像自動生成など飛躍的な性能向上を実現しています。最近では、OpenAI 社の GPT-4 などを筆頭に、莫大な知識と幅広い認識(マルチタスク²⁾)を実現する AI 基盤モデルの開発競争が世界的に行われています。しかし、これらの大規模 AI モデルは、性能向上に伴い、モデルサイズの増加(メモリー量の増加)と学習コストの上昇という課題を抱えており、開発には莫大な投資が必要になってきています。また、AI の適用先を産業への応用に限定した際には、必要とされる汎用性は必ずしも広範囲にわたるものではないため、限られた計算資源の中での軽量化が求められています。

本研究では、AI の適用先を適度に限定し、学習プロセスを最適化することで、軽量、低学習コストながら実用的な性能を達成する新しい AI 基盤モデルを開発しました。特に、特定のタスクへの適用(ファインチューニング)において、既存の大規模 AI モデルに匹敵する性能を達成しています。

【今後の展開】

現在、開発したモデルは単一画像入力のみ対応しておりますが、複数画像入力への対応を目指しています。これにより、鉄道車両の生産時に起こりうる説明性の高い異常検知への応用など、産業分野におけるさらなる応用の可能性が拓けます。

Press Release

【用語解説】

1) マルチモーダルモデル:

異なる種類のデータ(画像、テキスト、音声など)を統合して解析することができるモデル。
単一データを扱うモデルに比べ顕著に成果が出ている。

2) マルチタスクモデル:

複数の認識タスクを同時に学習することで、複数タスク(例えば画像認識、物体検出、画像の説明など)にまたがった知識を獲得するモデル。例えば、物体検出タスクには含まれない検出対象でも、画像の説明タスクで学習した知識を用いて物体検出が可能となる。

【論文情報】

雑誌名: IW-FCV 2024

論文タイトル: Constructing Lightweight Large Vision-Language Model

著者: 梁瀬 和哉、軸屋 敬介、表 英輝、土田 裕登、加藤 邦人(岐阜大学)

【問い合わせ先】

<研究に関すること>

東海国立大学機構 岐阜大学 工学部 教授 加藤邦人

電話: 058-293-2756

E-mail: kato.kunihito.k6@f.gifu-u.ac.jp

<報道に関すること>

東海国立大学機構 岐阜大学 総務部広報課広報グループ

電話: 058-293-3377

E-mail: kohositu@t.gifu-u.ac.jp