

## 研究課題別事後評価結果

1. 研究課題名： 与えられた指示文章に従い言語で判断を説明する AI

2. 個人研究者名

栗田 修平（理化学研究所 革新知能統合研究センター 研究員）

3. 事後評価結果

本研究は、AI が社会において人間と協働する際の重要な鍵となる、実世界における言語理解に焦点を当て、仮想世界や実世界と、言語情報を結びつけるベンチマークタスクとデータセットの構築の方法論と、実際のデータセットの構築を目的とした独創的かつ社会的な波及効果の高い研究である。本研究では、以下の二つの観点から研究を推進した。

(1) 実世界の情報と自然言語との対応付け技術については、テキストから参照された物体を画像や動画から探索する AI の基礎技術として「参照表現理解タスク」に着目し、主観視点動画からテキストで参照された物体を探すタスクや、全周カメラで撮影されたシーンをテキストにて説明するタスクなど、実世界とテキストをつなぐ各種のタスクを開発した。(2) 自然言語による人間の指示を理解し、実世界・仮想世界から情報を取得し、行動や質問応答を行う技術については、視覚と言語を融合した実世界探索において、画像キャプションモデルを利用して言語指示と動作・視覚の結合や、3次元の写実的なシーンからの質問応答、詳細な物体の探索などについて、タスクやモデルの研究開発と提案を行い、データセット構築を行なった。

これらの成果は、多数の優れた論文として、視覚・言語処理から深層学習にいたる AI 分野のトップ国際会議等に採択され、国際的にも注目されている。さらに、タスクとモデルの提案、データセットの公開などを通じて、今後の先端的な言語・視覚 AI の研究に不可欠な研究情報基盤の構築に貢献している点は国際的にも意義がある。これらは、従来の言語理解手法では扱うことができない物理世界の情報をテキストと対応付けて取扱うことを可能にする新しい「信頼される AI の基盤技術」の開発を促進・支援するものであり、学術・産業・社会的な波及効果が高い。今後の自然言語処理、画像処理、ロボティクスなどの幅広い分野の AI 技術基盤となると期待される。

本さきがけ研究を通じて、他分野のさきがけ研究者との連携や、実社会応用の検討などの活動も積極的に行い、AI 分野を横断した「実世界での言語理解」という新しい融合分野を着実に切り拓いている。また研究成果のオープンソース・オープンデータとしての国際的な公開による社会貢献も、今後の AI 研究の推進方法として高く評価できる。実世界で活躍できる AI システムの開発と評価に関する第一人者となることを期待する。実世界における言語理解の立場から、本研究で培った成果を、現在飛躍的に発展している生成モデル等の新しい AI システムで生じている種々の課題にも展開することを期待したい。