

信頼される AI の基盤技術  
2020 年度採択研究代表者

2022 年度  
年次報告書

栗田 修平

理化学研究所 革新知能統合研究センター  
特別研究員

与えられた指示文章に従い言語で判断を説明する AI

## 研究成果の概要

実世界での言語理解課題について、動画や 360 画像、3D のような多様な実世界の環境情報とテキスト情報を統合するために、参照表現理解や環境情報のテキスト説明をテーマとして、以下の3つの観点から研究を進めた。1つ目は、実世界を移動する人やロボットなどを念頭に、一人称視点動画から特定の物体をテキストに基づいて探索するための基盤整備である。既存の研究では、主として画像からテキストで指示された物体を探索していたが、実世界を移動するロボットなどに応用するには対象物体を含む適切な画像を入力として選択する必要があった。この問題点を指摘し、大規模な一人称視点動画セットからテキスト参照表現に基づいて特定の物体を探索するためのデータセットおよび参照表現理解手法と物体追跡手法を組み合わせた手法を提案した。2つ目は、実世界環境に近い写実的な仮想環境上で、与えられた質問に応じて環境内部をナビゲーションし、質問の正解物体を探して質問応答を行う研究を進めた。この研究は、大規模な3D シーンセット上で昨年度から進めていた質問応答アノテーションを活用し、国際電気通信基礎技術研究所(ATR)等の研究者と共同で行われた。また、3D シーンから日用品や調理器具などの小物を探索するためのデータセットを作成した。3つ目は360° カメラから得られる画像などマルチコンテキストな画像から、ユーザークエリに合わせて適切なキャプションを生成する研究である。3D 環境での質問応答データセット作成研究は、昨年度に続いて画像系のトップ国際会議である CVPR2022 で口頭発表を行った。意味役割ラベル付与において強化学習により文を処理する研究を言語系の国際会議 COLING2022 にて筆頭で発表を行った他、実世界における言語理解をキッチンドメインに応用するためのレシピコーパス整備研究について、京都大学およびオムロンの研究者と COLING2022 で発表を行うなど、基礎から応用を見据えた研究を進めている。

### 【代表的な原著論文情報】

- 1) Shuhei Kurita, Hiroki Ouchi, Kentaro Inui and Satoshi Sekine, Iterative Span Selection: Self-Emergence of Resolving Orders in Semantic Role Labeling, The 29th International Conference on Computational Linguistics (COLING2022) (October 2022).
- 2) Keisuke Shirai, Atsushi Hashimoto, Taichi Nishimura, Hiroataka Kameko, Shuhei Kurita, Yoshitaka Ushiku and Shinsuke Mori, Visual Recipe Flow: A Dataset for Learning Visual State Changes of Objects with Recipe Flows, The 29th International Conference on Computational Linguistics (COLING2022) (October 2022).
- 3) Daichi Azuma(\*), Taiki Miyanishi(\*), Shuhei Kurita(\*) and Motoaki Kawanabe, ScanQA: 3D Question Answering for Spatial Scene Understanding  
(\*): equally contributed  
The 2022 Conference on Computer Vision and Pattern Recognition (CVPR2022).
- 4) テキスト条件付き物体検出器と参照表現理解への応用  
MIRU2022, 姫路, (2022.7.25-28).  
桂尚輝, 栗田修平
- 5) QuIC-360: 複数のコンテキストを含む 360° 画像に対するクエリ指向画像説明文生成のため

のデータセット構築

言語処理学会第 29 回年次大会(NLP2023), 沖縄.

前田 航希, 栗田 修平, 宮西大樹

- 6) ARKitSceneRefer: 多様な実世界における 3D 屋内シーンでのテキストに基づいた小物の位置特定

言語処理学会第 29 回年次大会(NLP2023), 沖縄.

加藤 駿弥, 栗田 修平, Chenhui Chu, 黒橋 禎夫