

信頼される AI の基盤技術
2020 年度採択研究者

| |
|------------------|
| 2021 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

§ 1. 研究成果の概要

実世界での言語理解課題について、事前に収集された 3D 情報を活用して、実世界環境についての質問応答を行うテーマで研究を進めた。特に、本年度は、事前に収集された 3D 環境モデル上での質問応答データセット作成を中心に進めた。例えば, ”What is placed between the TV set and coffee table?” など, 実世界に紐付いた質問を理解し, 解答を探索する課題を考える。この課題では, 質問に登場する物体へナビゲーションすることだけではなく, ナビゲーション中に観測された画像情報などを統合して環境地図を構築し, 質問応答を行うことが求められる。実際に, ナビゲーションなどを通して実世界で収集・推定された複数枚の画像やカメラ位置情報は, SLAM や Photogrammetry のような手法で 3D 環境情報へと統合することができる。この研究では, このように質問応答に必要な環境情報が予め収集され 3D の環境モデルへと再構築されていることを前提に, 3D モデルから質問応答を行う 3D-QA 課題を提案し, データセットおよび 3D-QA モデルを作成した。この課題をナビゲーションモデルと組み合わせることで, ナビゲーションを通して 3D 地図を作成しながら, 質問応答を目指すモデルを作成することができる。この研究は京都大学の修士学生である東大地くん, および国際電気通信基礎技術研究所(ATR)の宮西大樹主任研究員, 川鍋一晃室長と共同で行われ, 画像系のトップ国際会議である CVPR2022 に採択され口頭発表を予定している(論文 1)。この他に, ニューヨーク大学の Kyunghyun Cho 先生との視覚と言語によるナビゲーションに関する国際共同研究について ICLR2021 にて発表を行った(論文 2)。また, 筑波大学の中山功太君との共著論文が EMNLP2021 Findings に採択された(論文 3)。加えて, 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共同で研究を行い, 自然言語処理年次大会で発表を行うなど, 実世界での言語理解課題研究は広がりを見せている。

【代表的な原著論文情報】

- 1) Daichi Azuma(*), Taiki Miyanishi(*), Shuhei Kurita(*) and Motoki Kawanabe, “ScanQA: 3D Question Answering for Spatial Scene Understanding” The 2022 Conference on Computer Vision and Pattern Recognition (CVPR2022). Oral. To appear.
(*): equally contributed.
- 2) Kouta Nakayama, Shuhei Kurita, Akio Kobayashi, Yukino Baba and Satoshi Sekine, “Co-Teaching Student-Model through Submission Results of Shared Task” The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP2021 Findings), (November 2021).
- 3) Shuhei Kurita and Kyunghyun Cho, “Generative Language-Grounded Policy in Vision-and-Language Navigation with Bayes’ Rule” Ninth International Conference on Learning Representations (ICLR2021), Online (May 2021).