

2023 年度年次報告書
信頼される AI の基盤技術
2021 年度採択研究代表者

松原 崇

大阪大学 大学院基礎工学研究科
准教授

望まれる性質を設計段階で保証する幾何学的深層学習の構築

研究成果の概要

これまで、物理現象を計測データから学習によってモデル化するという枠組みにおいて、そこに存在する対称性などの数理的構造を保存することで、質的に優れたモデルを設計することに注力してきた。本年はそのような手法やアイデアをコンピュータビジョンに応用する研究に注力した。

1)では、ある物体を撮影すると、その距離に応じて画像中のスケール(大きさ)が変わるという幾何学的な構造に注目した。通常の深層学習では、様々なスケールに対応できる多くのフィルタを用意する必要があり、必要学習データの増大、エッジケースへの汎化性の低下などが起こる。そこで、自動運転などで用いられるセンシング法である LiDAR に注目する。この手法では対象までの距離を計測するため、スケールを計算できる。そして、畳み込みニューラルネットワークのフィルタを偏微分作用素であるとみなし、スケールに合わせて補正することで、あらゆるスケールを単一のカーネルで処理できるように拡張した。結果的に物体までの距離に堅牢な認識が可能となった。

2)においては画像生成における言語の論理構造に注目した。拡散モデルを用いたテキストに基づく画像生成では、テキスト中の単語と画像中のピクセルの対応関係(アテンション)を取得し、それに基づいて画像を更新していく。しかし、テキストの一部や構造が往々にして無視されてしまう。例えば「黒い犬と白い猫」や「ボールを持った少年」の描かれた画像を生成しようとしているのに「黒い猫」や「ボールを蹴っている少年」が生成されたりする。これは文章が持つ論理構造が無視されているからであると考えた。そこでテキストの意味を明示的に存在量子や含意を用いた論理命題で表現し、さらにアテンションが原子論理式に対応するという仮定の下、論理命題が満たされるようにアテンションを操作する手法を開発した。これによって、テキストの意図に忠実な画像生成を実現した。

【代表的な原著論文情報】

- 1) Hidetaka Marumo and Takashi Matsubara, "Scale-equivariant convolution for semantic segmentation of depth image," Nonlinear Theory and Its Applications, IEICE, Jan. 2024.
- 2) Kota Sueyoshi and Takashi Matsubara, "Predicated Diffusion: Predicate Logic-Based Attention Guidance for Text-to-Image Diffusion Models," Proc. of The IEEE/CVF Computer Vision and Pattern Recognition Conference 2024 (CVPR2024), Jun. 2024.