

2023 年度
創発的研究支援事業 年次報告書

研究担当者	中島悠太
研究機関名	大阪大学
所属部署名	データビリティフロンティア機構
役職名	教授
研究課題名	映像記述のための言語を創出する人工知能の実現
研究実施期間	2023 年 4 月 1 日～2024 年 3 月 31 日

研究成果の概要

映像記述のための言語の創出に向け、その基礎的な技術となりうる対照学習を利用した画像記述の学習について研究開発を実施した。近年、同一の画像に対して各種画像処理を施し、ニューラルネットワークに入力してそれらが同一の画像であることを（他の画像と比較して）見分けることにより、ニューラルネットワークを学習する手法が事前学習として広く利用されている。このとき、画像中の対応する領域（画素）を見つけ出し、それらの表現が類似するように学習することで、画像レベルではなく、画像中のオブジェクトレベルで意味のある表現を学習できる。本研究では、画素のレベルでの対応を考えることにより、オブジェクトの境界をより先鋭化できることを示した。これは単語を利用して映像を記述するアプローチにとって、オブジェクトを考慮した言語創発の一助となることが期待される。本成果はコンピュータビジョン分野の国際会議である WACV2024 にて発表した。これに加えて、画像から言語を創発するモデルについて初期的検討を開始した。このモデルは、画像を入力すると言語的表現を出力するエンコーダと、言語的表現を入力するとそのベクトル表現を出力するデコーダから構成されており、画像から得られるベクトル表現と言語的表現を介して得られるベクトル表現が近くなるように学習する。これにより、画像内容を記述する言語的表現を学習により獲得する。MNIST データセットによる実験では、言語的表現からの識別で 90%以上の正答率が得られており、通常のベクトル表現による識別と比べて遜色のない性能が得られている。