

2022年 2月 24日

豊橋技術科学大学長 殿

情報・知能工学専攻
学位審査委員会
委員長 南 哲人

論文審査及び最終試験の結果報告

このことについて、博士学位論文審査を実施し、下記の結果を得ましたので報告いたします。

学位申請者	Rizal Setya Perdana		学籍番号	第 189303 号
申請学位	博士 (工学)	専攻名	大学院工学研究科博士後期課程 情報・知能工学 専攻	
博士学位 論文名	Contextualization of Multimodal Sequence Models on Image to Text in Story Generation (物語生成における画像からテキストへのマルチモーダルシーケンスモデルの文脈化)			
論文審査の 期間	2021年 7月 15日 ~ 2022年 2月 24日			
公開審査会 の日	2022年 2月 24日	最終試験の 実施日	2022年 2月 24日	
論文審査の 結果※	合格		最終試験の 結果※	合格
審査委員会(学位規程第6条)				
学位申請者にかかる博士学位論文について、論文審査、公開審査会及び最終試験を行い、別紙論文内容の要旨及び審査結果の要旨のとおり確認したので、学位審査委員会に報告します。				
委員長	三浦 純			
委員	秋葉 友良		石田 好輝	
		印		印
		印		印

※論文審査の結果及び最終試験の結果は「合格」又は「不合格」の評語で記入すること。

論文内容の要旨

画像シーンを理解し言葉で表現することは人間にとっては簡単であるが、機械にとっては困難なタスクである。近年、機械学習技術の発展により複雑な画像シーンを理解することが可能になってきており、特に画像から言語を生成する分野では、画像シーケンスから自然言語により物語生成する研究が行われている。本論文は、画像シーケンスから物語生成の際に文脈を利用すること（文脈化）の重要性を指摘し、物語生成における画像からテキストへの変換を行うための、画像とテキストを同時に扱うマルチモーダルシーケンスモデルを文脈化するための新たな表現についての研究をまとめたものである。

第1章では、研究の背景、対象とする問題の定義、焦点となる問題、研究目的と本論文の貢献、論文構成について述べている。第2章では、関連研究について基礎となる概念を整理して述べている。第3章では、インスタンススペースのマルチモーダル表現について関連研究と本論文で提案する表現を述べ、実験結果と評価について述べている。第4章ではマルチモーダルシーケンスについて、関連研究と本論文で提案する相関解析と注意制御メカニズムに基づくモデルについて述べ、実験結果と評価について述べている。第5章では、物語生成における画像からテキストへの変換に用いるマルチモーダルシーケンスモデルの、クロスモーダルな注意制御メカニズムと事前学習したテキスト生成モデルに基づく文脈化手法を提案し、実験によりその有効性を示している。第6章では、本論文の成果をまとめるとともに、今後の課題と展望を述べている。

審査結果の要旨

機械学習技術の実用化のためには、人間と同様に文脈を適切に利用することが必須である。本論文では、画像シーケンスから自然言語による物語生成を行う問題を対象とし、画像とテキストというマルチモーダルシーケンスを表現する新たなモデルを提案し、さらに相関解析と注意制御メカニズムを用いて文脈化することによる物語生成手法を提案しており、高い新規性を有する。本論文の主要な成果は以下のようにまとめられる。

- 1) 画像とテキスト間の意味的相関を、インスタンススペースではなくシーケンススペースで抽出することにより、言語生成において画像中に出現する対象の記述に加え、画像中に現れない非視覚的概念の単語を生成できることを示した。文中において、非視覚的概念の単語があることにより、生成された物語が単なる対象物の記述に比べてより理解しやすいものになることを示している。
- 2) 言語生成における文脈化の改良により、自然言語による物語生成の質を向上させることができることを示している。この文脈化のモデルは、文脈がない場合に、語彙サイズが小さく、単調な文章となることを改良することに焦点をあてた戦略を取り込んだものである。

これらの成果は学術的に高い評価を受けるとともに、機械学習関連分野への幅広い応用が期待される。以上より、本論文は博士（工学）の学位論文に相当するものと判定した。

(各要旨は1ページ以上可)