

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	Broadening the Context: User-specific Information and Visual Scenes for Conversation Tasks
著者(和文)	FIKRIABDURRISYAD
Author(English)	Abdurrisyad Fikri
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11757号, 授与年月日:2022年3月26日, 学位の種別:課程博士, 審査員:奥村 学,熊澤 逸夫,中山 実,篠崎 隆宏,船越 孝太郎
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11757号, Conferred date:2022/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名		Abdurrisyad Fikri	
		氏名	職名		氏名	職名
論文審査 審査員	主査	奥村学	教授	審査員	船越孝太郎	准教授
	審査員	熊澤 逸夫	教授			
		中山実	教授			
		篠崎隆宏	准教授			

### 論文審査の要旨 (2000 字程度)

本論文は「Broadening the Context: User-specific Information and Visual Scenes for Conversation Tasks」と題し、英文全5章より構成されている。

第1章「Introduction」では、本研究の背景を説明するとともに、本論文の目的を述べている。人間との対話を行う技術においては、対話の文脈を単に用いるだけではなく、それ以外の情報も文脈とともに活用する必要があることをまず述べている。そして、この課題を解決するため、本論文では、ユーザレベルの類似性および、画像と言語の両方を用いるマルチモーダルな設定を導入した研究に取り組むことを説明している。さらに、マルチモーダルな設定を導入した対話研究においては、対話シーンとその中の注目する部分および、対応する発話対からなる新しいマルチモーダル対話データセットを構築することを説明している。

第2章「Related Work」では、本研究で取り組む2つのタスク、対話における応答生成と、画像と言語の両方を用いるマルチモーダルな対話に関する先行研究について説明している。

第3章「User-specific Response Generation」では、ユーザ固有の情報を導入した応答生成の枠組みについて述べている。提案する応答生成の枠組みでは、ユーザ固有の情報をを用いることにより、そのユーザのスタイルに沿った応答を生成できること、訓練データ中に存在しない未知のユーザにも対応できることを説明している。そして、提案手法がベースライン手法と比べて性能を全体的に向上できること、具体的には、ベースライン手法より実際のユーザに似た応答を生成できることを実験的に示している。

第4章「Response Classification with Visual Scene Dataset」では、マルチモーダルな設定を導入した対話の枠組みについて述べている。まず、対話シーンと対応する発話対からなる、構築したマルチモーダル対話データセットについて説明している。構築したデータセットは、第3者視点でのデータセットとしてこれまでにない特徴を有しており、対話を行う複数の人間のそばに位置し、その対話を第三者視点で理解する対話エージェント構築に有用であることを説明している。また、このデータセットを分析することで、どのような対話シーン画像やその中の注目する部分が対話に有効であるかを明らかにしている。そして、このデータセット上のマルチモーダルな対話タスクのためのニューラルネットワークに基づくモデルを提案し、最後に、提案手法がベースライン手法と比べて性能を全体的に向上できることを実験的に示している。

第5章「Conclusion and Future Work」では、本研究の結論と今後の課題を説明している。今後の課題では残された重要な課題を述べている。

以上を要するに、本論文は、人間との対話を行う技術において重要な文脈以外の情報として、ユーザ固有の情報を導入した応答生成のためのニューラルネットワークに基づくモデルを提案するとともに、これまでにない第3者視点でのマルチモーダル対話データセットを構築した上で、それらの有効性を検証しており、工学上貢献するところが大きい。よって博士(工学)の学位を授与するに十分な価値を持つものと認められる。

注意:「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。