JAIST Repository
https://dspace.jaist.ac.jp/

Title	Conversational Context-Aware Physiological and Linguistic Fusion for Self-reported Sentiment Analysis
Author(s)	Duc, Tran Minh
Citation	
Issue Date	2023-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18736
Rights	
Description	Supervisor: 岡田 将吾, 先端科学技術研究科, 修士(情報科学)

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Master's Thesis

Conversational Context-Aware Physiological and Linguistic Fusion for Self-reported Sentiment Analysis

TRAN MINH DUC

Supervisor: Prof. SHOGO OKADA

Graduate School of Advanced Science and Technology
Japan Advanced Institute of Science and Technology
(Information Science)

September, 2023

Abstract

Developing dialog systems capable of dynamically adapting to a user's sentiment state in real time is a challenging task. Existing multimodal models have demonstrated impressive performance in estimating third-party labeled sentiment levels by incorporating features from linguistic, visual, and speech modalities. On the other hand, physiological signals play a crucial role in estimating self-reported sentiment as they exhibit involuntary changes associated with emotions. Previous studies have shown the effectiveness of fusing physiological and linguistic features for self-reported sentiment estimation.

However, these studies often neglect the contextual interaction between exchanges, where each exchange consists of a pair of system and user utterances. To address this gap, we propose an efficient approach that incorporates interplay of physiological features between exchanges in dialogue, which consists of a system utterance followed by a user utterance. Specifically, we introduce a framework that combines linguistic and physiological signals across exchanges. Our approach employs attention mechanisms to capture contextual information and long-term dependencies within the dialog, enabling a comprehensive understanding of sentiment evolution. Additionally, we leverage convolutional neural networks (CNNs) to learn robust representations from physiological signals, enhancing the interpretation of the user's emotional changes.

Through extensive experiments, our approach surpasses existing multimodal models in sentiment estimation. Our findings highlight the importance of inter-exchange learning for effective sentiment adaptation in dialog systems. By considering time-series changes in linguistic and physiological features across multiple exchanges, our approach captures the dynamical changes of sentiment level, leading to more accurate and adaptive dialog interactions.