Paper ID: 3642

DUALGATS: DUAL GRAPH ATTENTION NETWORKS FOR EMOTION RECOGNITION IN CONVERSATIONS

Duzhen Zhang, Feilong Chen, Xiuyi Chen
Baidu Inc & Huawei Inc, Beijing, P.R.China
zhangduzhen@baidu.com, chenfeilong10@huawei.com, chenxiuyi01@baidu.com

Abstraction

Capturing complex contextual dependencies plays a vital role in Emotion Recognition in Conversations (ERC). Previous studies have predominantly focused on speaker-aware context modeling, overlooking the discourse structure of the conversation. In this paper, we introduce Dual Graph ATtention networks (DualGATs) to concurrently consider the complementary aspects of discourse structure and speaker-aware context, aiming for more precise ERC. Specifically, we devise a Discourse-aware GAT (DisGAT) module to incorporate discourse structural information by analyzing the discourse dependencies between utterances. Additionally, we develop a Speaker-aware GAT (SpkGAT) module to incorporate speaker-aware contextual information by considering the speaker dependencies between utterances. Furthermore, we design an interaction module that facilitates the integration of the DisGAT and SpkGAT modules, enabling the effective interchange of relevant information between the two modules. We extensively evaluate our method on four datasets, and experimental results demonstrate that our proposed DualGATs surpass state-of-the-art baselines on the majority of the datasets.

Index Terms— Emotion Recognition in Conversations, Graph Attention Networks, Sentiment Analysis

Model Overview

In this paper, we propose a novel method called Dual Graph ATtention networks (DualGATs) that aims to improve the accuracy of ERC by simultaneously considering the complementarity of discourse structure and speaker-aware context. The DualGATs layer comprises three components: Discourse-aware GAT (DisGAT), Speaker-aware GAT (SpkGAT), and an interaction module with a differential regularizer. The overall architecture of our DualGATs is illustrated in Figure 1.

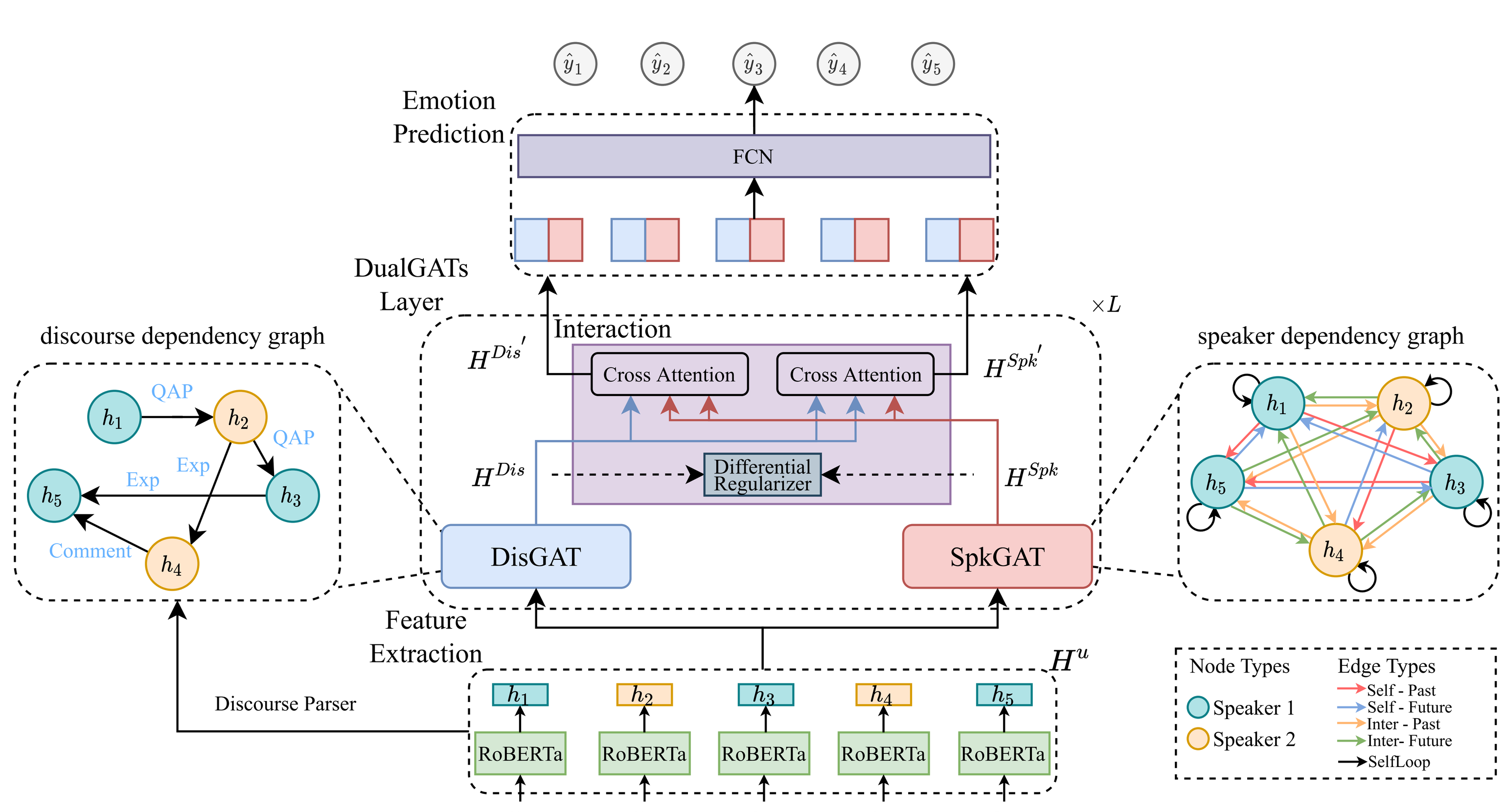

Figura 1: The overall architecture of our DualGATs, encompassing three essential modules: DisGAT, SpkGAT, and Interaction. DisGAT propagates discourse structural information by leveraging discourse dependencies between utterances, while SpkGAT propagates speaker-aware contextual information considering speaker and temporal dependencies. The interaction module initially employs a differential regularizer to ensure that the DisGAT and SpkGAT modules capture distinct contextual information. Subsequently, it utilizes mutual cross-attention to integrate the DisGAT and SpkGAT modules, facilitating the exchange of relevant information between them. In the diagram, the discourse dependency types *Question-Answer Pair* (QAP) and *Explanation* (Exp) are denoted.

Our contributions can be summarized as follows:

- We propose DualGATs to simultaneously consider the complementarity of discourse structure and speaker-aware context for more precise and accurate ERC.
- We introduce an interaction module to exchange the relevant information between the SpkGAT and DisGAT modules by mutual cross-attention, where a differential regularizer is proposed to induce the two modules to capture different contextual information.
- We conduct extensive experiments on four publicly available ERC datasets. The results of our experiments demonstrate that DualGATs outperform state-of-the-art baselines on most of the tested datasets. Further analyses validate the effectiveness of the critical components in DualGATs.

Datasets

We evaluate our DualGATs on the following four ERC datasets. The statistics of these four datasets are drawn in Table 1.

Tabela 1: The statistics of four ERC datasets.

Dataset	# Conversations			# Uterrances		
	Train	Val	Test	Train	Val	Test
IEMOCAP	120		31	5810		1623
MELD	1038	114	280	9989	1109	2610
EmoryNLP	659	89	79	7551	954	984
DailyDialog	11118	1000	1000	87170	8069	7740

Experimental results

The overall performance of all the compared baselines and our DualGATs on the four datasets is reported in Table 2. The results of our experiments demonstrate that DualGATs outperform state-of-the-art baselines on most of the tested datasets.

Tabela 2: The overall performance of all the compared baselines and our DualGATs on four ERC datasets. Bold font denotes the best performance. The marker $*$ refers to significant test $p\text{-}value < 0.05$ comparing with CoMPM, the marker † refers to significant test $p\text{-}value < 0.05$ comparing with CoG-BART, and the marker ‡ refers to significant test $p\text{-}value < 0.05$ comparing with DialogueRole. Moreover, we refer to the results from DialogueRole with the marker ♣, from DAG-ERC with the marker ♠, from SGED with the marker ♢, and the results for the remaining baselines are from original papers.

Models	IEMOCAP	MELD	EmoryNLP	DailyDialog
BC-LSTM♣	54.95	56.87	-	50.24
ICON♣	58.54	-	-	-
DialogueRNN♠	62.75	57.03	-	-
+RoBERTa♠	64.76	63.61	37.44	57.32
DialogueCRN	66.20	58.39	-	-
+RoBERTa♢	66.46	63.42	38.91	-
KET	59.56	58.18	34.39	53.37
DialogueGCN	64.18	58.10	-	-
+RoBERTa♠	64.91	63.02	38.10	57.52
RGAT	65.22	60.91	34.42	54.31
+RoBERTa♠	66.36	62.80	37.89	59.02
DialogXL	65.94	62.41	34.73	54.93
DAG-ERC	68.03	63.65	39.02	59.33
CoG-BART	66.18	64.81	39.04	56.29
CoMPM	66.33	66.52	37.37	60.34
COSMIC	65.28	65.21	38.11	58.48
TODKAT[3]	61.33	65.47	38.69	58.47
DialogueRole	**68.23**	65.34	-	60.95
CauAIN	67.61	65.46	-	58.21
DisGCN	64.10	64.22	36.38	-
DualGATs (Ours)	67.68	**66.90***	**40.69**†	**61.84**‡

Experiment Analysis

We perform ablation studies to analyze the effects of critical modules in our DualGATs, shown in Table 3. Overall, our DualGATs with all modules achieve the best performance.

Tabela 3: Experimental results of ablation study.

Models	IEMOCAP	MELD	EmoryNLP	DailyDialog
DisGAT	64.56	64.23	37.65	58.96
SpkGAT	66.32	64.66	38.34	59.91
DualGATs w/o regularizer	66.70	65.73	39.53	60.93
DualGATs w/o cross attention	66.43	65.46	39.68	60.26
DualGATs (Ours)	**67.68**	**66.90**	**40.69**	**61.84**