

GRAPHEME: GRAPH NEURAL NETWORKS WITH MULTI-EXPERT FUSION FOR EMOTION-CAUSE PAIR EXTRACTION

Xue Gu^{1,2}, Ziyao Meng^{1,2}, Adriano Tavares¹, Tiago Gomes¹, Hao Xu^{2,*}

¹University of Minho, Department of Industrial Electronics, Guimarães, Portugal

²Jilin University, College of Computer Science and Technology, Changchun, China

ABSTRACT

Emotion-Cause Pair Extraction (ECPE) aims to identify clause-level pairs expressing emotions and their underlying causes. Existing methods rely on monolithic architectures that model causal relations through homogeneous cues, such as syntactic or sequential proximity, making them fragile to long-range or semantically complex associations. We argue that emotion-cause reasoning is inherently heterogeneous, emerging from the confluence of diverse linguistic signals including syntactic, semantic role, coreference, temporal, and commonsense. To capture this complexity, we propose **GraphEME**, a multi-expert graph neural network that decomposes inference into specialized linguistic perspectives, each encoded as a subgraph and processed independently. A learnable gating mechanism dynamically fuses expert outputs, enhancing both robustness and interpretability. Experiments on Chinese ECPE benchmarks show that **GraphEME** achieves state-of-the-art performance, with ablation and case studies validating the necessity and effectiveness of multi-expert fusion.

Index Terms— Emotion-Cause Pair Extraction, Multi-expert, Graph Neural Network

1. INTRODUCTION

Emotion-Cause Pair Extraction (ECPE) aims to identify pairs of emotion clauses and their corresponding cause clauses from the document. Unlike coarse-grained sentiment analysis, ECPE requires detecting emotion and cause events, modeling cross-clausal semantic dependencies, and relying on implicit commonsense reasoning and diverse linguistic cues. Understanding why an emotion occurs, rather than just recognizing it, makes ECPE essential for applications such as psychological modeling and dialogue systems.

Early approaches adopt two-step pipelines that first extract emotion and cause clauses separately before pairing them [1]. While intuitive, these methods suffer from cascade error. To mitigate this, end-to-end models based on sequence labeling or graph neural networks (GNNs) have been proposed, formulating ECPE as a joint tagging or clause-pair classification task [2, 3, 4, 5, 6, 7]. However, these methods rely heavily on local structure, limiting their capacity to capture long-range or semantically complex associations. More recently, machine reading comprehension (MRC) and question-answering (QA)-inspired models reformulate ECPE as a span prediction problem conditioned on emotion queries [8, 9, 10, 11],

improving context sensitivity. With the rise of large language models (LLMs), prompt-based methods further enable ECPE without explicit structural modeling [12, 13]. Despite these advances, most existing methods integrate diverse signals, such as semantics, temporality, or commonsense knowledge, within a monolithic architecture that applies a uniform inference mechanism across all instances. This homogeneous integration limits their capacity to preserve the semantic integrity of individual cues and adaptively weight their contributions based on context.

Inspired by cognitive science, where humans consult multiple experts before complex judgments. We propose **GraphEME**, a multi-expert GNN framework that decomposes ECPE inference into multiple linguistic views (e.g., syntactic, temporal, commonsense). Each view encoded as a subgraph over clauses with domain-specific edges. A dedicated encoder processes each subgraph, preserving cue-specific semantics, and a learnable gating mechanism adaptively fuses the outputs based on context.

The key contributions of this work are as follows:

- We introduce a cognitive-inspired paradigm by reframing ECPE as a **multi-expert collaborative reasoning** process, where heterogeneous linguistic cues are explicitly disentangled into specialized reasoning pathways, offering a more principled alternative to monolithic modeling.
- We propose **GraphEME**, a GNN framework that instantiates this paradigm through dedicated subgraphs for syntactic, temporal, semantic, and commonsense dependencies, fused via a learnable gating mechanism that enables both robustness and interpretability.
- Experimental results show that our method achieves state-of-the-art performance.

2. METHOD

The proposed **GraphEME** framework, illustrated in Figure 1, consists of four sequential components: (1) a clause-level context encoder, (2) multi-view graph construction, (3) multi-expert graph encoding and fusion, and (4) emotion-cause pair prediction. We formalize ECPE as a collaborative process across specialized linguistic experts, each responsible for capturing distinct types of causal evidence. This modular architecture enables the explicit disentanglement of heterogeneous signals, such as syntax, semantics, and commonsense, and their context-aware integration, enhancing both model robustness and interpretability.

2.1. Clause-level Context Encoder

Given a document with n clauses: $D = \{c_1, c_2, \dots, c_n\}$, we encode each clause independently using a shared BERT encoder [14].

*Corresponding author. Email: xuhao@jlu.edu.cn

This work is supported by the National Natural Science Foundation of China (NSFC, No. 62476111), the Department of Science and Technology of Jilin Province (No. 20230201086GX), the Science and Technology Bureau of Changchun (No. 24GNYZ11), and the Industry University Research Innovation Fund of the Ministry of Education project (No. 2022XF017).

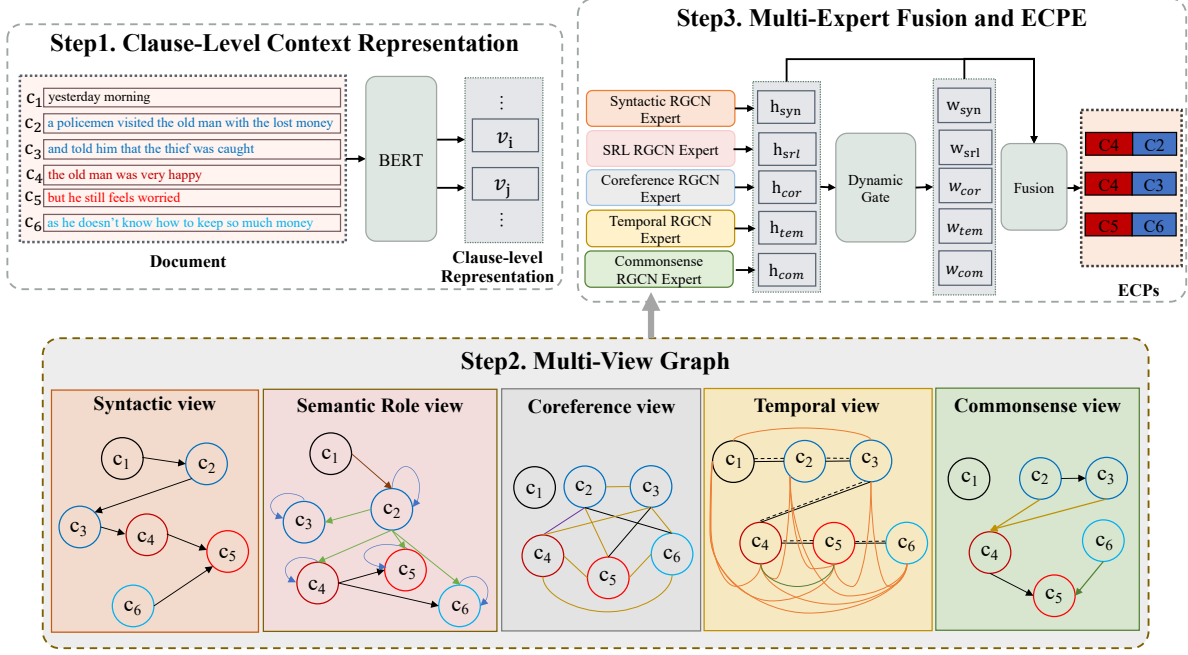


Fig. 1. Overview of the proposed multi-view multi-expert framework for ECPE.

For clause c_i , we form the input sequence: $\mathbf{x}_i = [\text{CLS}] c_i [\text{SEP}]$, truncated or padded to a fixed length L . The encoded token representations are: $\mathbf{H}_i = \text{BERT}(\mathbf{x}_i) \in \mathbb{R}^{L \times d}$, where d is the hidden size. We extract the final hidden state of the $[\text{CLS}]$ token as the clause-level embedding: $\mathbf{v}_i = \mathbf{H}_i[0, :]$, $i = 1, \dots, n$. The set of initial node embeddings is denoted as $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \in \mathbb{R}^{n \times d}$, which serve as inputs to all subsequent graph views.

2.2. Multi-View Graph Construction

Emotion-cause relations are inherently *multi-faceted*, emerging from the interplay of diverse linguistic phenomena, including syntactic dependency, semantic roles, temporal order, referential coherence, and commonsense knowledge. To explicitly model this heterogeneity, we decompose the relational structure of the document into five complementary graph views, each grounded in a principled linguistic framework and designed to capture a specific type of causal cue.

Formally, given a document with n clauses, we construct a multi-view graph ensemble:

$$\mathcal{G} = \{\mathcal{G}^{(k)} = (\mathcal{V}, \mathcal{E}^{(k)}) \mid k \in \mathcal{K}\} \quad (1)$$

where $\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is the shared node set, $\mathbf{v}_i$ representing a clause representation, and $\mathcal{E}^{(k)}$ is the edge set for view k , and $\mathcal{K} = \{\text{syn}, \text{srl}, \text{cor}, \text{tem}, \text{com}\}$ denotes the five linguistic views: syntactic, semantic role, coreference, temporal, and commonsense views, respectively.

Syntactic View ($\mathcal{G}^{\text{syn}}$): Clause-Level Dependency Aggregation. We construct the syntactic view graph $\mathcal{G}^{\text{syn}} = (\mathcal{V}, \mathcal{E}^{\text{syn}})$ to model structural dependencies between clauses. Using LTP [15] for dependency parsing, we derive clause-level edges from word-level syntactic relations. Let $A_w \in \{0, 1\}^{N_w \times N_w}$ be the directed adjacency matrix of the dependency tree, and $R \in \{0, 1\}^{N_c \times N_w}$ the clause-to-word incidence matrix. We compute the clause-aggregated

dependency strength via:

$$M_{\text{clause}} = R A_w R^\top \quad (2)$$

where $M_{\text{clause}}[i, j]$ reflects the total number of syntactic dependencies between clauses which are calculated by the word dependency. We binarize M_{clause} to obtain a directed edge set $\mathcal{E}^{\text{syn}}$. This yields a sparse, syntactically grounded graph that supports local and cross-clausal dependency modeling.

Semantic Role View ($\mathcal{G}^{\text{srl}}$): Predicate-Argument Alignment. We construct $\mathcal{G}^{\text{srl}} = (\mathcal{V}, \mathcal{E}^{\text{srl}}, \mathbf{A}^{\text{srl}})$ to capture cross-clausal argument-to-predicate alignments grounded in semantic role labeling. Unlike syntactic dependencies, semantic roles encode deep semantic relations between events and their participants, such as agency, causation, and temporality, providing direct cues for ECPE. Using an off-the-shelf SRL parser (e.g., LTP [15]), we extract predicates and their labeled arguments for each clause. For each argument a_k with role r_k in clause C_i , if its head noun or canonical form matches a mention in another clause C_j ($j \neq i$), we add a directed edge $(j, i) \in \mathcal{E}^{\text{srl}}$. When available, we also incorporate coreference links from a separate module (see §3.2.3). Each edge is annotated with the role label r_k in the edge attribute matrix $\mathbf{A}^{\text{srl}}$.

This construction captures long-range semantic dependencies, particularly causal traces (ARGM-CAU) that are highly indicative of valid ECPE. By grounding edges in linguistically motivated roles, $\mathcal{G}^{\text{srl}}$ introduces a structured inductive bias for relational reasoning, complementing other views that rely on syntax or commonsense.

Coreference View ($\mathcal{G}^{\text{cor}}$): Referential Coherence. We construct $\mathcal{G}^{\text{cor}} = (\mathcal{V}, \mathcal{E}^{\text{cor}}, \mathbf{A}^{\text{cor}})$ to model referential coherence. Coreference chains are identified using the LTP toolkit [15], which provides rule-based and neural coreference resolution for Chinese texts. For each chain χ , let $\mathcal{I}_\chi \subseteq \{1, \dots, n\}$ denote the involved clause indices. We add undirected edges between all distinct pairs: $(i, j) \in \mathcal{E}^{\text{cor}}$ for $i, j \in \mathcal{I}_\chi, i \neq j$. Edge attributes in $\mathbf{A}^{\text{cor}}$ encode mention types (PRONOUN.REF, ZERO.REF, etc.).

This approach enables robust modeling of long-range referential links, particularly those involving implicit agents that are critical for accurate emotion-cause alignment.

Temporal View ($\mathcal{G}^{\text{tem}}$): Event Ordering. We construct $\mathcal{G}^{\text{tem}} = (\mathcal{V}, \mathcal{E}^{\text{tem}}, \mathbf{A}^{\text{tem}})$ to model fine-grained temporal relations. Directed NEXT edges $(i, i + 1)$ connect adjacent clauses. REALIZATION edges link prior events to clauses containing cognitive verbs (e.g., “realized”), detected via a predefined verb list. CAUSE_HINT edges connect cause to emotion clauses when triggered by conjunctions (e.g., “because”, “since”). Undirected PARALLEL edges join clauses marked as simultaneous (e.g., “at that time”). All triggers are identified through pattern matching with domain-adapted rules.

Edge types in $\mathbf{A}^{\text{tem}}$ enable the GNN to distinguish narrative order from event-time precedence, resolving ambiguity in emotion-cause alignment.

Commonsense View ($\mathcal{G}^{\text{com}}$): Commonsense Knowledge Injection. We construct $\mathcal{G}^{\text{com}} = (\mathcal{V}, \mathcal{E}^{\text{com}}, \mathbf{A}^{\text{com}})$ to inject structured commonsense knowledge into ECPE. Edges are derived from the Chinese ATOMIC [16], which models inferential chains between events and their emotional or intentional consequences. We extract four types of directed edges: CAUSE→EMOTION links causal events to emotional responses; INTENT→EMOTION connects goal-oriented clauses to emotional outcomes; EMOTION→REACTION captures behavior driven by emotions; and SELF_CAUSE marks self-contained emotion elicitation. Built from external knowledge, this view supports reasoning over implicit emotion-cause dynamics through semantic plausibility.

2.3. Multi-Expert Graph Encoder and Fusion

To capture heterogeneous linguistic signals from multiple views (syntax, SRL, coreference, temporal, commonsense), we construct a relational graph. For each view $k \in \mathcal{K}$, we apply a 2-layer RGCN [17] to encode the subgraph. At layer l , the update for node i is:

$$\mathbf{h}_i^{(l,k)} = \sigma \left(\sum_{r \in \mathcal{R}_k} \sum_{j \in \mathcal{N}_r(i)} \frac{1}{|\mathcal{N}_r(i)|} \mathbf{W}_r^{(l,k)} \mathbf{h}_j^{(l-1,k)} \right) \quad (3)$$

where $\mathcal{R}_k$ is the relation set in view k , and $\mathcal{N}_r(i)$ denotes neighbors of i under relation r .

We adopt a node-wise gating mechanism to integrate expert outputs. Our gate is computed *after* multi-view reasoning, enabling context-aware weighting. Specifically, we first apply an initial fusion to obtain a preliminary fused representation:

$$\mathbf{z}_i = \text{LayerNorm} \left(\text{MLP}_{\text{fusion}}^{(1)}([\mathbf{h}_i^{(1,L)}; \dots; \mathbf{h}_i^{(5,L)}]) \right) \quad (4)$$

where $\mathbf{h}_i^{(k,L)}$ is the output of the final RGCN layer in view k . This intermediate representation captures cross-view interactions and serves as input to the gating network:

$$\mathbf{g}_i = \text{softmax}(\text{MLP}_{\text{gate}}(\mathbf{z}_i)) \in \mathbb{R}^5 \quad (5)$$

Each expert’s output is then reweighted: $\tilde{\mathbf{h}}_i^{(k)} = g_{i,k} \cdot \mathbf{h}_i^{(k,L)}$. The final fused embedding is obtained through a second fusion layer:

$$\mathbf{h}_i^{\text{fused}} = \text{LayerNorm} \left(\text{MLP}_{\text{fusion}}^{(2)}([\tilde{\mathbf{h}}_i^{(1)}; \dots; \tilde{\mathbf{h}}_i^{(5)}]) \right) \quad (6)$$

This two-stage design allows the model to dynamically adjust the contribution of each expert based on the outcomes of multi-view reasoning, enhancing both adaptivity and interpretability. The gating weights $\mathbf{g}_i$ can be visualized to reveal which linguistic cues dominate for each clause.

2.4. Emotion-Cause Pair Prediction

Given fused clause embeddings $\{\mathbf{h}_i^{\text{fused}}\}_{i=1}^n$, we compute pairwise scores:

$$s(i, j) = \mathbf{w}^\top \text{MLP}(\text{concat}(\mathbf{h}_i^{\text{fused}}, \mathbf{h}_j^{\text{fused}}, |\mathbf{h}_i^{\text{fused}} - \mathbf{h}_j^{\text{fused}}|)) \quad (7)$$

with probabilities $p_{ij} = \sigma(s(i, j))$.

During inference, we select all pairs (i, j) for which $p_{ij} > 0.5$. If no pair exceeds this threshold, we default to the pair with the highest score to ensure non-empty output. The model is trained end-to-end with binary cross-entropy:

$$\mathcal{L} = - \sum_{(i,j)} [y_{ij} \log p_{ij} + (1 - y_{ij}) \log(1 - p_{ij})] \quad (8)$$

with negative sampling for balance. The gating weights $\mathbf{g}_i$ further provide interpretability by highlighting view-specific contributions to each clause.

3. EXPERIMENTS

3.1. Dataset and Metrics

We conduct experiments on the benchmark ECPE dataset introduced by Xia and Ding [1]. The dataset contains 1,945 news documents from SINA city news. Following the standard setting [1], we adopt 10-fold cross-validation, repeat experiments 10 times, and report the average Precision (P), Recall (R), and F1-score (F1) on the main task of ECPE, as well as its two subtasks: emotion clause extraction (EE) and cause clause extraction (CE).

3.2. Results and Analysis

Comparison with SOTA models. Table 1 compares **GraphEME** against SOTA methods on the ECPE benchmark dataset. Our model achieves the highest F1-score of **81.69**, outperforming the previous state-of-the-art method MM-R (80.62) by +1.07 points, while also attaining the best precision (**83.17**) and competitive recall (80.25). This demonstrates that decomposing linguistic reasoning into specialized graph-based experts effectively captures complex emotion-cause dependencies. On the subtasks, **GraphEME** achieves strong results: 93.65 F1 on EE, only outperformed by MM-R’s 97.38 but with higher precision (92.42 vs. 90.38), and **84.79** F1 on CE, surpassing all baselines. Notably, several models exhibit imbalanced performance, e.g., Guided-QA excels in CE (87.60 R) but lags in EE (72.90 R), while MM-R dominates EE but underperforms in CE. In contrast, **GraphEME** delivers consistently high performance across all components, indicating its balanced capability in modeling both emotional expressions and causal triggers through multi-view structural reasoning.

Ablation Study. To evaluate the contribution of each linguistic view and the fusion mechanism, we conduct ablation experiments on **GraphEME**. As shown in Table 2, removing any single expert leads to performance degradation, confirming the importance of multi-view modeling. The largest drop occurs when the commonsense module is removed ($\Delta\text{F1} = -2.33$), highlighting its critical role in capturing implicit causal relations. Disabling the temporal view results in a -1.81 F1 drop, indicating that temporal order is vital for cause-effect alignment. Syntax and SRL views also contribute meaningfully (-0.94 and -0.76 F1), especially for local dependency and predicate-argument structure modeling.

Replacing the learnable gating mechanism with uniform averaging reduces F1 by -1.84 , demonstrating that adaptive weighting is

Table 1. Performance comparison (%) on the ECPE dataset. Best results in **bold**.

Model	ECPE			EE			CE		
	P	R	F1	P	R	F1	P	R	F1
Indep-EC (2019) [1]	67.21	57.05	61.28	83.64	81.07	82.30	70.41	60.83	65.07
A ² Net (2022) [18]	75.03	77.80	76.34	90.67	90.98	90.80	77.62	79.20	78.35
UECA-Prompt (2022) [19]	71.82	77.99	74.70	84.75	91.95	88.16	76.24	79.16	77.55
PBJE (2022) [6]	79.22	73.84	76.37	90.77	86.91	88.76	81.79	76.09	78.78
MM-R (2022) [10]	82.18	79.27	80.62	97.38	90.38	93.70	83.28	79.64	81.35
MMN (2023) [20]	76.11	73.96	75.02	90.37	87.85	89.07	79.01	75.54	77.21
Guided-QA (2023) [8]	71.90	79.20	75.40	77.10	69.20	72.90	84.70	90.80	87.60
KMGP (2024) [21]	74.25	62.08	67.32	85.22	84.43	84.79	72.75	61.33	66.42
GAT-ECPE (2024) [22]	72.65	77.52	74.92	90.98	91.03	90.99	76.17	78.72	77.34
SEG (2025) [23]	77.74	78.44	78.06	92.10	90.40	91.22	80.22	79.85	80.00
DGRE (2025) [24]	72.20	77.48	74.69	90.96	91.54	91.22	75.14	78.55	76.75
GraphEME (Ours)	83.17	80.25	81.69	93.65	92.42	93.05	84.79	80.78	82.75

Table 2. Ablation study on GraphEME (%).

Configuration	P	R	F1
Full Model	83.17	80.25	81.69
w/o Syntax	82.10	79.40	80.75
w/o SRL	82.35	78.90	80.93
w/o Temporal	81.80	78.10	79.88
w/o Commonsense	80.90	77.80	79.36
w/o Gating (Mean Fusion)	81.20	78.50	79.85
Single Unified Graph	78.00	75.10	76.52

essential for balancing heterogeneous signals across views. When all experts are replaced with a single unified graph encoder, performance drops sharply by -5.17 F1, underscoring the benefit of decomposition and specialization. These results validate the effectiveness of our multi-expert design and context-aware fusion strategy.

Case Study & Visualization. To elucidate the inner workings of GraphEME, we present a detailed case study in Figure 2, illustrating how multi-view linguistic reasoning, particularly commonsense knowledge, enables accurate ECPE in the presence of implicit causality. In this case: the emotional response in (c5) is not triggered by an overtly negative event, nor is the cause clause syntactically adjacent. Instead, the model recognizes that going under a bridge, especially one in a remote, uninhabited area, is an unusual choice for a routine outing. This atypical behavior, when combined with the lack of public facilities, suggests a potential underlying concern, a pattern consistent with commonsense knowledge about human behavior (e.g., in ATOMIC).

As shown in the Figure 2 (c), the commonsense expert assigns a dominant weight of 0.75, far exceeding other views. This distribution indicates that the attribution relies primarily on abductive inference over commonsense knowledge, rather than surface-level cues. The low syntactic weight reflects the absence of causal connectives; the temporal view captures proximity but lacks semantic depth.

The Figure 2 (b) visualizes the inferred reasoning path: clause (c3) (“going under the bridge”) and context (c6) (“no facilities”) jointly activate latent intents such as $x\text{Intent}$: “to escape reality” or “to seek solitude”. These trigger the observer reaction $o\text{React}$: “feels surprise”, aligning with empathic responses to perceived behavioral anomalies.

Notably, the absence of explicit affective language or temporal markers underscores the necessity of externalized knowledge integration. GraphEME’s modular architecture, guided by view-

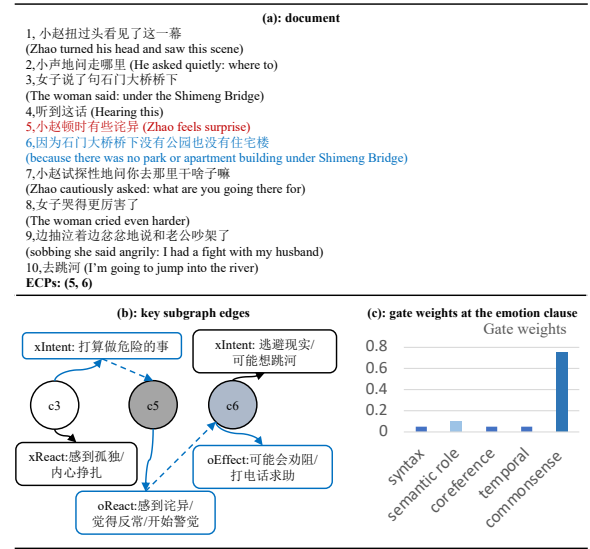


Fig. 2. Case study showing how GraphEME uses commonsense knowledge to infer implicit ECPs. (a): input document; (b): inferred reasoning path with ATOMIC-based concepts; (c): learned gate weights.

specific gate weights, effectively prioritizes the most relevant evidence, enabling detection of emotionally salient events before full situational disclosure.

4. CONCLUSION

We present GraphEME, a multi-expert graph neural network framework for ECPE that reframes the task as a collaborative reasoning process over heterogeneous linguistic cues. Unlike existing methods that rely on monolithic modeling, GraphEME explicitly disentangles diverse signals, including syntactic, semantic role, coreference, temporal, and commonsense knowledge, into specialized subgraphs, each processed by a dedicated encoder. A learnable gating mechanism dynamically fuses these views, enabling both robust prediction and interpretable reasoning. Extensive experiments on benchmark datasets demonstrate that GraphEME outperforms SOTA methods. In the future, we plan to extend this framework to dialogic settings.

5. REFERENCES

- [1] Rui Xia and Zixiang Ding, "Emotion-cause pair extraction: A new task to emotion analysis in texts," in *Proceedings of the ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*. 2019, pp. 1003–1012, Association for Computational Linguistics.
- [2] Chuang Fan, Chaofa Yuan, Jiachen Du, Lin Gui, Min Yang, and Ruifeng Xu, "Transition-based directed graph construction for emotion-cause pair extraction," in *Proceedings of the ACL 2020, Online, July 5-10, 2020*. 2020, pp. 3707–3717, Association for Computational Linguistics.
- [3] Chaofa Yuan, Chuang Fan, Jianzhu Bao, and Ruifeng Xu, "Emotion-cause pair extraction as sequence labeling based on A novel tagging scheme," in *Proceedings of the EMNLP 2020, Online, November 16-20, 2020*. 2020, pp. 3568–3573, Association for Computational Linguistics.
- [4] Zifeng Cheng, Zhiwei Jiang, Yafeng Yin, Na Li, and Qing Gu, "A unified target-oriented sequence-to-sequence model for emotion-cause pair extraction," *IEEE ACM Transactions on Audio, Speech and Language Processing*, vol. 29, pp. 2779–2791, 2021.
- [5] Ying Chen, Wenjun Hou, Shoushan Li, Caicong Wu, and Xiaoqiang Zhang, "End-to-end emotion-cause pair extraction with graph convolutional network," in *Proceedings of the COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*. 2020, pp. 198–207, International Committee on Computational Linguistics.
- [6] Junlong Liu, Xichen Shang, and Qianli Ma, "Pair-based joint encoding with relational graph convolutional networks for emotion-cause pair extraction," in *Proceedings of the EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*. 2022, pp. 5339–5351, Association for Computational Linguistics.
- [7] Siyuan Chen and Kezhi Mao, "A graph attention network utilizing multi-granular information for emotion-cause pair extraction," *Neurocomputing*, vol. 543, pp. 126252, 2023.
- [8] Huu-Hiep Nguyen and Minh-Tien Nguyen, "Emotion-cause pair extraction as question answering," in *Proceedings of the ICAART 2023, Volume 3, Lisbon, Portugal, February 22-24, 2023*. 2023, pp. 988–995, SCITEPRESS.
- [9] Ting-Wei Chang, Yao-Chung Fan, and Arbee L. P. Chen, "Emotion-cause pair extraction based on machine reading comprehension model," *Multimedia Tools and Applications*, vol. 81, no. 28, pp. 40653–40673, 2022.
- [10] Changzhi Zhou, Dandan Song, Jing Xu, and Zhijing Wu, "A multi-turn machine reading comprehension framework with rethink mechanism for emotion-cause pair extraction," in *Proceedings of the COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*. 2022, pp. 6726–6735, International Committee on Computational Linguistics.
- [11] Zifeng Cheng, Zhiwei Jiang, Yafeng Yin, Cong Wang, Shiping Ge, and Qing Gu, "A consistent dual-mrc framework for emotion-cause pair extraction," *ACM Transactions on Information Systems*, vol. 41, no. 4, pp. 105:1–105:27, 2023.
- [12] Zengzhi Wang, Qiming Xie, Zixiang Ding, Yi Feng, and Rui Xia, "Is chatgpt a good sentiment analyzer? A preliminary study," 2023.
- [13] Jialiang Wu, Yi Shen, Ziheng Zhang, and Longjun Cai, "Enhancing large language model with decomposed reasoning for emotion cause pair extraction," 2024.
- [14] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1*, Jill Burstein, Christy Doran, and Thamar Solorio, Eds. 2019, pp. 4171–4186, Association for Computational Linguistics.
- [15] Wanxiang Che, Yunlong Feng, Libo Qin, and Ting Liu, "N-LTP: A open-source neural chinese language technology platform with pretrained models," *CoRR*, vol. abs/2009.11616, 2020.
- [16] Dawei Li, Yanran Li, Jiayi Zhang, Ke Li, Chen Wei, Jianwei Cui, and Bin Wang, "C³KG: A Chinese commonsense conversation knowledge graph," in *Findings of the Association for Computational Linguistics: ACL 2022*, Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, Eds., Dublin, Ireland, May 2022, pp. 1369–1383, Association for Computational Linguistics.
- [17] Michael Sejr Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling, "Modeling relational data with graph convolutional networks," in *The Semantic Web - 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, June 3-7, 2018, Proceedings*. 2018, vol. 10843 of *Lecture Notes in Computer Science*, pp. 593–607, Springer.
- [18] Shunjie Chen, Xiaochuan Shi, Jingye Li, Shengqiong Wu, Hao Fei, Fei Li, and Donghong Ji, "Joint alignment of multi-task feature and label spaces for emotion cause pair extraction," in *Proceedings of the COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*. 2022, pp. 6955–6965, International Committee on Computational Linguistics.
- [19] Xiaopeng Zheng, Zhiyue Liu, Zizhen Zhang, Zhaoyang Wang, and Jiahai Wang, "Ueca-prompt: Universal prompt for emotion cause analysis," in *Proceedings of the COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*. 2022, pp. 7031–7041, International Committee on Computational Linguistics.
- [20] Xichen Shang, Chuxin Chen, Zipeng Chen, and Qianli Ma, "Modularized mutuality network for emotion-cause pair extraction," *IEEE ACM Transactions on Audio, Speech and Language Processing*, vol. 31, pp. 539–549, 2023.
- [21] Linlin Zong, Jinglin Zhang, Jiahui Zhou, Xianchao Zhang, and Bo Xu, "Emotion-cause pair extraction via knowledge-driven multi-classification and graph-based position embedding," *Appl. Intell.*, vol. 54, no. 3, pp. 2703–2715, 2024.
- [22] Peican Zhu, Botao Wang, Keke Tang, Haifeng Zhang, Xiaodong Cui, and Zhen Wang, "A knowledge-guided graph attention network for emotion-cause pair extraction," *Knowl. Based Syst.*, vol. 286, pp. 111342, 2024.
- [23] Yuwei Wang, Yuling Li, Kui Yu, and Jing Yang, "A semantic structure-based emotion-guided model for emotion-cause pair extraction," *Pattern Recognit.*, vol. 161, pp. 111296, 2025.
- [24] Ming Zhang, Zhe Chen, Vasile Palade, Tao Lu, Liya Wang, Junchi Zhang, and Yanduo Zhang, "Double-graph representation with relational enhancement for emotion-cause pair extraction," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 36, no. 6, pp. 10859–10873, 2025.