Refine, Align, and Aggregate: Multi-view Linguistic Features Enhancement for Aspect Sentiment Triplet Extraction

Guixin Su[1,2,3,4], Mingmin Wu[1,2,3,4], Zhongqiang Huang[1,2,3,4], Yongcheng Zhang[1], Tongguan Wang[1,2,3,4], Yuxue Hu[1,2,3,4], Ying Sha[1,2,3,4]*

[1]College of Informatics, Huazhong Agricultural University, Wuhan, China
[2]Key Laboratory of Smart Farming for Agricultural Animals, Wuhan, China
[3]Hubei Engineering Technology Research Center of Agricultural Big Data, Wuhan, China
[4]Engineering Research Center of Intelligent Technology for Agriculture, Ministry of Education

{cometsue,wmm_nlp,hzq,zhyc,wang_tg}@webmail.hzau.edu.cn, {hyx,shaying}@mail.hzau.edu.cn

Abstract

Aspect Sentiment Triplet Extraction (ASTE) aims to extract the triplets of aspect terms, their associated sentiments and opinion terms. Previous works based on different modeling paradigms have achieved promising results. However, these methods struggle to comprehensively explore the various specific relations between sentiment elements in multi-view linguistic features, which is the prior indication effect for facilitating sentiment triplets extraction, requiring to align and aggregate them to capture the complementary higher-order interactions. In this paper, we propose ***M**ulti-**v**iew **L**inguistic **F**eatures **E**nhancement* (MvLFE) to explore the aforementioned prior indication effect in the "*Refine, Align, and Aggregate*" learning process. Specifically, we first introduce the *relational graph attention network* to encode the word-pair relations represented by each linguistic feature and refine them to pay more attention to the aspect-opinion pairs. Next, we employ the *multi-view contrastive learning* to align them at a fine-grained level in the contextual semantic space to maintain semantic consistency. Finally, we utilize the *multi-semantic cross attention* to capture and aggregate the complementary higher-order interactions between diverse linguistic features to enhance the aspect-opinion relations. Experimental results on several benchmark datasets show the effectiveness and robustness of our model, which achieves state-of-the-art performance.

1 Introduction

Aspect Sentiment Triplet Extraction (ASTE) is a new variant of *Aspect-Based Sentiment Analysis* (ABSA) (Pontiki et al., 2014), which is an information extraction style task to identify all sentiment triplets from a review to explain WHAT the targeted aspects are, HOW their sentiment polarities are and WHY they have such polarities (i.e., opinion terms). A sentence with triplets marking in Figure 1 illustrates the definition of the ASTE task.

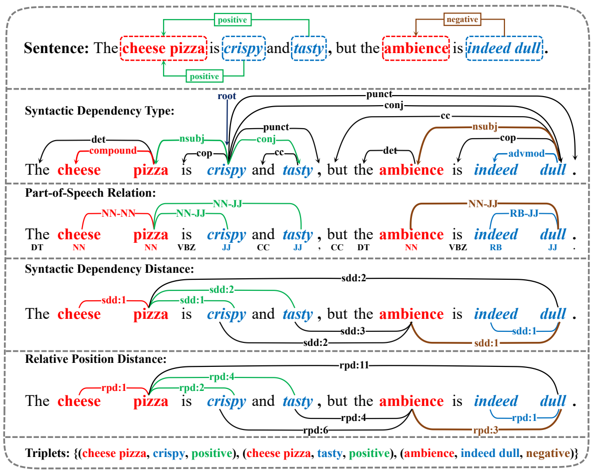

Figure 1: The prior indication effect of multi-view linguistic features for ASTE. It reveals various specific relations between matched aspect and opinion terms, also within the terms themselves in different views of linguistic features, which facilitate the extraction of sentiment triplets from input sentences. In the above triplet set, aspect and opinion terms are marked in red and blue, with *positive* sentiment in green and *negative* in brown. "*sdd/rpd:n*" denotes distance between two words is *n*.

Early work adopts the straightforward solution to decompose the ASTE task into several ABSA subtasks to separately extract sentiment elements of triplets with a two-stage pipeline approach (Peng et al., 2020). To exploit the association among multiple subtasks, Mao et al. (2021); Chen et al. (2021a); Liu et al. (2022) transform the ASTE task into the machine reading comprehension (MRC) paradigm with diverse task-specific queries, which jointly train multiple subtasks but pipeline inference. Obviously, these methods potentially lead to the well-known error propagation problem.

To alleviate this problem, most of recent studies focus on jointly extracting the sentiment triplets in an end-to-end fashion (Xu et al., 2020, 2021; Chen et al., 2022b; Yu et al., 2023). Some efforts formulate the ASTE task as a text generation problem

*Corresponding author.

Findings of the Association for Computational Linguistics ACL 2024, pages 3212–3228
August 11-16, 2024 ©2024 Association for Computational Linguistics

to generate the sentiment triplets in one shot (Yan et al., 2021; Gao et al., 2022; Gou et al., 2023). However, these approaches are constrained in capturing reciprocity among the sentiment elements.

Therefore, Jing et al. (2021); Zhang et al. (2022); Liang et al. (2023) construct the word-pair relations by designing a novel grid tagging scheme (Wu et al., 2020a) to capture such association. Chen et al. (2021b); Fei et al. (2022); Chen et al. (2022a) further leverage the syntax dependency or additional linguistic features to enrich the word-pair representations. Nevertheless, these methods primarily fuse the linguistic features and the contextual features at a shallow level, which neglect to align and aggregate them to capture the complementary higher-order interactions. Consequently, they fail to exploit the prior indication effect of multi-view linguistic features for ASTE.

Naturally, two questions arise regarding the prior indication effect by our observations.

Q1: What is the prior indication effect for ASTE? We interpret it as there exist various specific relations between matched aspect-opinion pairs, also within aspect or opinion terms themselves in different views of linguistic features, which facilitate the sentiment triplets extraction. Take Figure 1 as an example. First, in the view of **syntactic dependency type**, we observe that the aspect term is the nominal subject of the opinion term due to the *"nsubj"* type. Besides, *"cheese"* and *"pizza"* comprise an aspect term by the *"compound"* type, while *"indeed"* is the adverbial modifier of *"dull"* so they combine as an opinion term. Thus, these dependency types facilitate not only the extraction of aspect and opinion terms but also their pairing. Second, in the view of **part-of-speech relation**, we find that aspect terms are nouns while opinion terms are adjectives. Hence, the word pair with the *"NN-JJ"* part-of-speech combination tends to form an aspect-opinion pair. In addition, we notice that the matched aspect and opinion terms are closer in terms of the **distance of syntactic dependency and relative position**.

Q2: How to effectively exploit the prior indication effect to help ASTE? We argue that exploring it requires the "*Refine, Align, and Aggregate*" learning process. First, there exist numerous irrelevant word-pair relations represented by multi-view linguistic features to ASTE (e.g., relations between commas and other words). Therefore, we need to refine the multi-view linguistic features to focus on the relations of the matched aspect-opinion pairs. Second, different views of linguistic features represent various relation types for the same aspect-opinion pairs, so they require to be aligned to maintain semantic consistency. Finally, since the prior indication effect of a single linguistic feature is one-sided for ASTE, it is crucial to capture and aggregate the complementary higher-order interactions between diverse linguistic features.

Motivated by the above observations, we propose ***M**ulti-**v**iew **L**inguistic **F**eatures **E**nhancement* (MvLFE) to explore the prior indication effect to enhance the word-pair relations. Firstly, we introduce the *relational graph attention network* to encode the word-pair relations represented by each linguistic feature as multi-channel edge features, which are normalized as doubly stochastic matrices to facilitate the aggregation and updating of word nodes, preventing over-smoothing across multi-layer learning. Moreover, edge features are updated by the edge refining strategy which considers the implicit results of aspect and opinion terms extraction to support their matching. Secondly, inspired by Image-Text Matching for cross-modal alignment, we employ the *multi-view contrastive learning* to align diverse aspect-opinion relations represented by different linguistic features in the semantic space which obtained by the biaffine transformation of contexts to maintain semantic consistency. Finally, we utilize the *multi-semantic cross attention* to aggregate the complementary higher-order interactions between diverse linguistic features to enhance aspect-opinion relations.

In summary, the key contributions are as follows:

- We propose a novel MvLFE model that effectively exploits the prior indication effect of multi-view linguistic features in the "*Refine, Align, and Aggregate*" learning process.
- We present the *relational graph attention network* to refine the relation-aware word-pair representations of multi-view linguistic features to attend the aspect-opinion pairs.
- We employ the *multi-view contrastive learning* and the *multi-semantic cross attention* to align and aggregate diverse linguistic features to enhance the aspect-opinion relations.

2 Related Works

Aspect-Based Sentiment Analysis (ABSA) (Pontiki et al., 2014) has received wide attention in recent years. Early works focus on identifying the

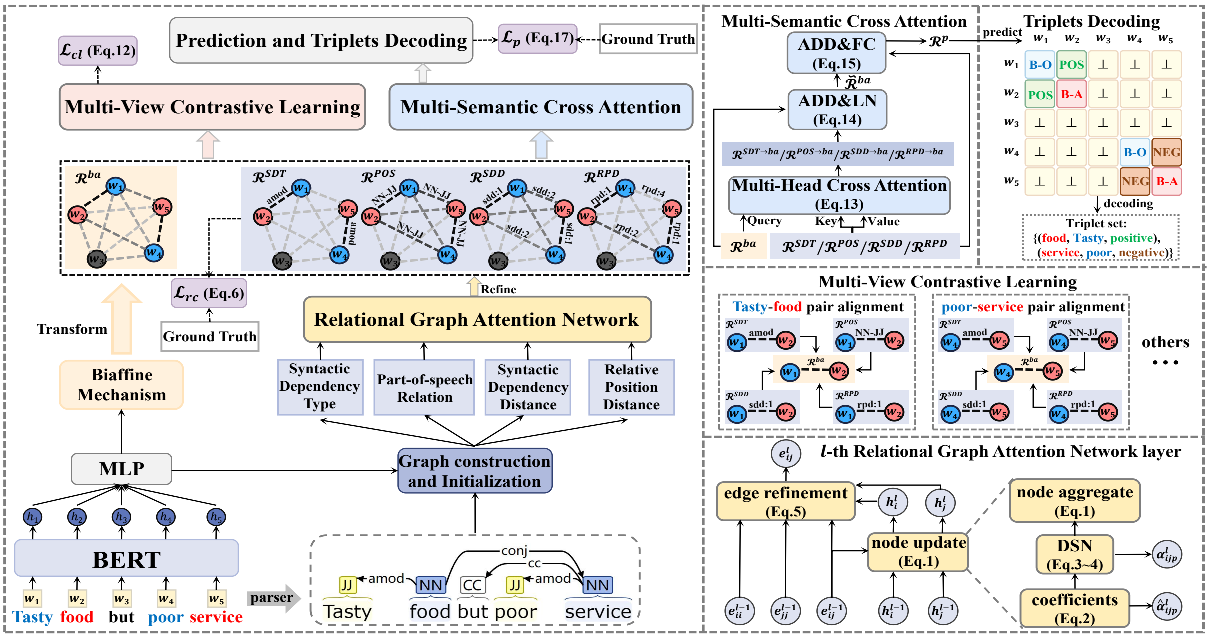

Figure 2: The overall architecture of our MvLFE model.

single sentiment element, i.e., *Aspect Term Extraction* (ATE) (Ma et al., 2019; Li et al., 2020; Wang et al., 2021), *Opinion Term Extraction* (OTE) (Dai and Song, 2019; Fan et al., 2019; Wu et al., 2020b) and *Aspect Sentiment Classification* (ASC) (Tang et al., 2022; Ma et al., 2023; Zhang et al., 2023). To further explore the interactions among sentiment elements, some efforts are devoted to coupling the individual subtasks, i.e., *Aspect-Opinion Pair Extraction* (AOPE) (Zhao et al., 2020; Chen et al., 2020b; Gao et al., 2021) and *Aspect Extraction and Sentiment Classification* (AESC) (He et al., 2019; Chen and Qian, 2020).

Based on their works, Peng et al. (2020) introduces the more challenging ASTE task to present a near-complete solution for ABSA. Subsequent works address ASTE task by different modeling paradigms: Sequence tagging (Xu et al., 2020; Zhang et al., 2020), Grid tagging (also known as table-filling) (Wu et al., 2020a; Chen et al., 2021b; Jing et al., 2021; Fei et al., 2022; Chen et al., 2022a; Zhao et al., 2022; Zhang et al., 2022; Liang et al., 2023), MRC-based methods with diverse task-specific queries (Mao et al., 2021; Chen et al., 2021a; Liu et al., 2022; Zhai et al., 2022) and Generative methods (Yan et al., 2021; Zhang et al., 2021; Gao et al., 2022; Mao et al., 2022; Lv et al., 2023; Mukherjee et al., 2023; Gou et al., 2023). However, these approaches generally ignore the prior indication effect of multi-view linguistic features for ASTE, which facilitates the triplet extraction.

3 Methodology

In this section, we first introduce the task definition and the adopted tagging scheme. Then we elaborate on the details of our MvLFE model and the training objective. The overall architecture is illustrated in Figure 2.

3.1 Task Definition and Tagging Scheme

Given an input sentence $X = \{w_1, w_2, \cdots, w_n\}$ with n words, the objective of the ASTE task is to extract a sentiment triplets set $\mathcal{T} = \{(a, o, s)_k\}_{k=1}^{|\mathcal{T}|}$ from sentence X, where triplet (a, o, s) refers to (*aspect term, opinion term, sentiment polarity*) and $s \in \{Positive, Neutral, Negative\}$.

Following by Grid Tagging Scheme (Wu et al., 2020a), we formulate the ASTE task as a unified tagging task by labeling the relations of all word pairs. Take Figure 3 as an example, the left side shows the tagging results for the sentence, while the right side explains the meaning of all tags of word-pair relations we used for ASTE. Specifically, we utilize tags {B-A, I-A, B-O, I-O} to determine the beginning and inside of aspect and opinion terms in the main diagonal of the grid. And we employ tags {A, O} to combine two words that belong to the same aspect term or opinion term in the non-diagonal region of the grid. Meanwhile, the sentiment tags {POS, NEU, NEG} are used not only to match the valid aspect-opinion pairs but also to judge their sentiment polarities.

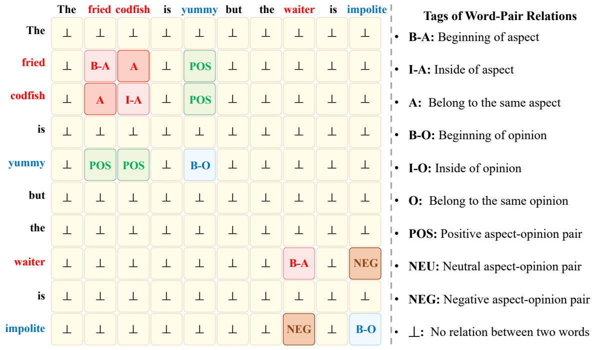

Figure 3: Grid tagging for triplet extraction in a sentence is illustrated, with each cell representing a word-pair relation. The meanings of all tags of word-pair relations are detailed on the right side.

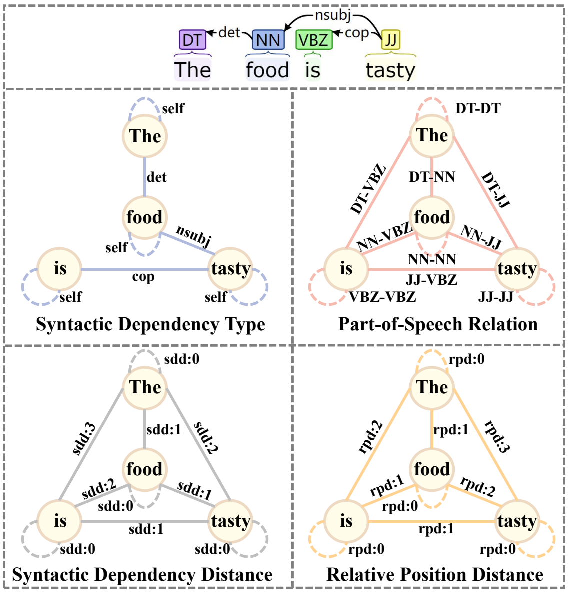

Figure 4: Word-pair relation graph construction of multi-view linguistic features. In each graph, the dashed edge is the self-loop for a word, and the solid edge is the relation between two words. "*sdd/rpd:n*" denotes the relational edge with distance *n* between two words.

3.2 Input and Encoding Layer

Pre-trained models have demonstrated their effectiveness in Natural Language Understanding (NLU) tasks (Qiu et al., 2020). In this work we utilize BERT (Devlin et al., 2019) as the context encoder. For a given sentence $X = \{w_1, w_2, \cdots, w_n\}$ with n words, the encoding layer outputs the hidden contextual representations $\mathcal{H} = \{h_1, h_2, \cdots, h_n\}$ at the last Transformer block.

3.3 Relational Graph Attention Network

3.3.1 Word-Pair Relation Graph Construction

Multi-view linguistic features of a sentence manifest various relations of each word pair. Therefore, it is intuitive to construct each view of linguistic feature as a word-pair relation graph.

Specifically, we utilize off-the-shelf CoreNLP toolkit (Manning et al., 2014) to parse sentences to obtain the different views of linguistic features for word-pair relations, including syntactic dependency type, part-of-speech relation, syntactic dependency distance and relative position distance. For each linguistic feature, we construct it as an undirected heterogeneous graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V}$ and $\mathcal{E}$ are sets of nodes and edges. Each word in the sentence is regarded as a node, while word-pair relations represented by the linguistic feature are considered edges. Figure 4 illustrates an example of the word-pair relation graph construction for multi-view linguistic features. For syntactic dependency type, we add a "*self*" dependency type to represent self-loop edge for each word node. For part-of-speech relation, we combine the part-of-speech tags as word-pair relation between two words. Besides, we further consider the distances between two words within the syntactic dependency tree and the sentence sequence as word-pair relations respectively. We initialize the node features with the hidden contextual representations of words and maintain different trainable embedding look-up tables to represent the edge features of various word-pair relation graphs.

3.3.2 Linguistic Features Refinement

Next, we introduce *relational graph attention network* (RGAT) to refine the word-pair relations represented by each linguistic feature, denoise the irrelevant word-pair relations to the ASTE and pay more attention to the valid aspect-opinion pairs.

Specifically, for each word-pair relation graph constructed by diverse linguistic features, we denote the node features of w_i and w_j as $h_i \in \mathbb{R}^{d_h}$ and $h_j \in \mathbb{R}^{d_h}$ respectively, where d_h is the dimension of hidden contextual representation. And $e_{ij} \in \mathbb{R}^P$ is defined as the relational edge feature represented by the word-pair (w_i,w_j), where P signifies the channels of edge feature to quantify the association degree of the word-pair relation to each pre-defined tag. Features of each node will be updated by aggregating the features of neighboring nodes and simultaneously incorporating the corresponding edge features. The operation at the l-th RGAT layer is defined as follows:

$$h_i^l = \mathop{\Big\|}_{p=1}^{P} \sigma\Big(\sum\nolimits_{j=1}^{n} \alpha_{ijp}^l W^l h_j^{l-1}\Big) \qquad (1)$$

where $\|$ denotes the concatenation operator, σ is the ELU activation function, and W^l is the weight matrix for the linear transformation of the inputs at the l-th layer. Notably, α_{ijp}^l denotes the attention coefficients between word nodes h_i^{l-1} and h_j^{l-1} on the p-th channel of edge features. For a specific channel of edge features, the attention function is calculated by:

$$\hat{\alpha}_{ijp}^l = \exp\left\{L(a^T[W^l h_i^{l-1}||W^l h_j^{l-1}])\right\} e_{ijp}^{l-1} \quad (2)$$

$$\alpha_{ijp}^l = \mathrm{DSN}(\hat{\alpha}_{ijp}^l) \quad (3)$$

where L is the LeakyReLU activation function, and $a^T \in \mathbb{R}^{2d_h}$ denotes the weight vector. DSN is the doubly stochastic normalization formulated as:

$$\tilde{\alpha}_{ijp}^l = \frac{\hat{\alpha}_{ijp}^l}{\Sigma_{t=1}^n \hat{\alpha}_{itp}^l}, \quad \alpha_{ijp}^l = \sum_{t=1}^n \frac{\tilde{\alpha}_{itp}^l \tilde{\alpha}_{jtp}^l}{\Sigma_{s=1}^n \tilde{\alpha}_{stp}^l} \quad (4)$$

Unlike the row normalization used in vanilla graph attention network (Velickovic et al., 2018), DSN normalizes the attention coefficients guided by edge features into the doubly stochastic matrices, which possess properties of symmetric, positive semi-definite and the largest eigenvalue 1. These properties contribute to preventing the edge matrix from exploding or shrinking to zero to stabilize the diffusion process, also mitigating the over-smoothing problem across multi-layer RGAT learning (Wang et al., 2018; Chen et al., 2020a).

Subsequently, we update edge features using an edge refining strategy, leveraging the implicit results of aspect and opinion term extraction to refine word-pair relation representations. For example, if w_i is an aspect term and w_j an opinion term, their pair (w_i,w_j) is more likely to denote a sentiment relation. The strategy is represented as:

$$e_{ij}^l = f(e_{ij}^{l-1} \oplus e_{ii}^{l-1} \oplus e_{jj}^{l-1} \oplus h_i^l \oplus h_j^l) \quad (5)$$

where f is the multi-layer perceptron with softplus activation function, and $\oplus$ denotes concatenation.

3.3.3 Relation Constraint

In order to more precisely steer RGAT to refine the word-pair relations, at the last layer, we conduct constraint on the refined word-pair relations $\mathcal{R}^v = \{r_{11}^v, r_{12}^v, \cdots, r_{nn}^v\}$ represented by each linguistic feature:

$$\mathcal{L}_{rc} = -\sum_{v\in V}\sum_{i}^{n}\sum_{j}^{n}\sum_{c\in\mathcal{C}} \mathbb{I}(y_{ij} = c)\log(r_{ij|c}^v) \quad (6)$$

where V denotes multi-view linguistic features, $\mathbb{I}(\cdot)$ is the indicator function, y_{ij} is the ground truth of word pair (w_i,w_j), and $\mathcal{C}$ signifies pre-defined tags.

3.4 Multi-View Contrastive Learning

Although different views of linguistic features represent various relation types for the same word pairs, they typically harbor analogous semantic implications, such as jointly reflecting the matching of aspect and opinion terms or the composition of sentiment terms. To ensure that multi-view linguistic features complement and reinforce each other at semantic level, it is vital to align them into the shared semantic space to maintain semantic consistency.

Specifically, we utilize the biaffine mechanism to transform the hidden contextual representations $\mathcal{H}$ to construct the semantic space $\mathcal{R}^{ba} = \{r_{11}^{ba}, r_{12}^{ba}, \cdots, r_{nn}^{ba}\}$ for word-pair relations, due to its proven expressive power in modeling the complex semantics between words (Dozat and Manning, 2017). The process is formulated as:

$$r_{ij}^{ba} = h_i^{\mathrm{T}} U_1 h_j + (h_i \oplus h_j)^{\mathrm{T}} U_2 + b \quad (7)$$

where $r_{ij}^{ba} \in \mathbb{R}^P$ denotes semantic features of the word pair (w_i,w_j) with P dimensions. U_1, U_2 and b are trainable weights and bias.

Inspired by Image-Text Matching for aligning text words and image regions at a fine-grained level (Lee et al., 2018; Diao et al., 2021; Pan et al., 2023), we project each linguistic feature into the semantic space and align them at the level of word-pair relations. First, we compute the cosine similarity matrix for all word pairs between each linguistic feature $\mathcal{R}^v$ and semantic feature $\mathcal{R}^{ba}$, i.e.,

$$s_{ijk}^v = \frac{(r_{ij}^v)^T r_{ik}^{ba}}{||r_{ij}^v||||r_{ik}^{ba}||}, \quad i, j, k \in [1, n] \quad (8)$$

where s_{ijk}^v denotes the similarity between word pair (w_i,w_j) in linguistic feature and word pair (w_i,w_k) in semantic feature. Then we integrate the attended semantic feature to obtain the projection of the linguistic feature in the semantic space:

$$a_{ij}^v = \sum_{k=1}^n \frac{exp(s_{ijk}^v/\tau)}{\sum_{m=1}^n exp(s_{ijm}^v/\tau)} r_{ik}^{ba} \quad (9)$$

where τ is the temperature hyperparameter.

Subsequently, we further calculate the cosine similarity between each word pair relation represented by the linguistic feature and its projection in the semantic space to measure the degree of alignment between them, i.e.,

$$\mathcal{Q}(r_{ij}^v, a_{ij}^v) = \frac{(r_{ij}^v)^T a_{ij}^v}{||r_{ij}^v||||a_{ij}^v||} \quad (10)$$

Finally, we evaluate the overall degree of alignment of the linguistic feature in the semantic space by LogSumExp pooling, i.e.,

$$\mathcal{S}(\mathcal{R}^v, \mathcal{R}^{ba}) = \frac{1}{\lambda} \log \sum_{i=1}^{n} \exp(\lambda \mathcal{Q}(r_{ij}^v, a_{ij}^v)) \quad (11)$$

where λ is the scaling factor.

We align multi-view linguistic features into the shared semantic space as described above and minimize the hinge-based triplet loss with margin γ as the alignment objective, i.e.,

$$\mathcal{L}_{cl} = \sum_{v \in V} [\gamma - \mathcal{S}(\mathcal{R}^v, \mathcal{R}^{ba}) + \mathcal{S}(\mathcal{R}^v, \hat{\mathcal{R}}^{ba})]_+ \quad (12)$$

where $[x]_+ \equiv max(x, 0)$. We select the linguistic feature $\mathcal{R}^v$ and semantic feature $\mathcal{R}^{ba}$ of the same sentence as the positive pair, and take the semantic feature $\hat{\mathcal{R}}^{ba}$ of other sentences within a mini-batch as negative pairs with linguistic feature $\mathcal{R}^v$. In this study, we focus on the hardest negative pair in a mini-batch following by Faghri et al. (2017), which is given by $\hat{\mathcal{R}}^{ba} = argmax_{d \neq \mathcal{R}^{ba}} \mathcal{S}(\mathcal{R}^v, d)$.

3.5 Multi-Semantic Cross Attention

Since the prior indication effect of a single linguistic feature is one-sided for ASTE, such as the noun-adjective part-of-speech combination commonly observed in matched or invalid aspect-opinion pairs, it is necessary to capture and aggregate the complementary higher-order interactions between diverse linguistic features to comprehensively enhance representations of word-pair relations.

Specifically, we employ multi-head cross attention (Vaswani et al., 2017) to capture the unique semantic representations of each linguistic feature. For the h-th attention head, we take $\mathcal{R}_h^Q = W_h^Q \mathcal{R}^{ba}$, $\mathcal{R}_h^K = W_h^K \mathcal{R}^v$ and $\mathcal{R}_h^{Va} = W_h^{Va} \mathcal{R}^v$ as query, key and value. The process is defined as:

$$\mathcal{R}^{v \to ba} = \big\Vert_{h=1}^{H} softmax\left(\frac{\mathcal{R}_h^Q (\mathcal{R}_h^K)^T}{\sqrt{d_k}}\right) \mathcal{R}_h^{Va} \quad (13)$$

where d_k is the dimension of key. Then we adopt residual connection and layer normalization to aggregate the semantic feature $\mathcal{R}^{ba}$ and each view's unique semantic representations $\mathcal{R}^{v \to ba}$, i.e.,

$$\tilde{\mathcal{R}}^{ba} = LayerNorm(\mathcal{R}^{ba} + \sum_{v \in V} \mathcal{R}^{v \to ba}) \quad (14)$$

Finally, we obtain the enhanced representations of word-pair relations $\mathcal{R}^p = \{r_{11}^p, r_{12}^p, \cdots, r_{nn}^p\}$ as logits for prediction by:

$$\mathcal{R}^p = f_p(\tilde{\mathcal{R}}^{ba} + \sum_{v \in V} \mathcal{R}^v) \quad (15)$$

where f_p is the fully connected network.

3.6 Training Objective

Our objective is to minimize the training loss as:

$$\mathcal{L} = \mathcal{L}_p + \beta \mathcal{L}_{rc} + \mu \mathcal{L}_{cl} \quad (16)$$

where hyperparameters β and μ are for adjusting the impact of corresponding relation constraint loss and contrastive learning loss. The standard cross-entropy loss $\mathcal{L}_p$ is used for the ASTE task, i.e.,

$$\mathcal{L}_p = -\sum_{i}^{n} \sum_{j}^{n} \sum_{c \in \mathcal{C}} \mathbb{I}(y_{ij} = c) \log(r_{ij|c}^p) \quad (17)$$

The triplets decoding is detailed in Appendix A.1.

4 Experiments

4.1 Datasets

We evaluate our proposed MvLFE model on four benchmark ASTE datasets (**ASTE-Data-V2**[1]) released by Xu et al. (2020), which is a refined version of the **ASTE-Data-V1** (Peng et al., 2020) to consider cases where one opinion is associated with multiple aspects. Those datasets are derived from Pontiki et al. (2014, 2015, 2016) with one in the laptop domain and three in the restaurant domain. Detailed statistics are shown in Appendix A.2.

4.2 Baselines

We compare our MvLFE model with the following four types of previous state-of-the-art methods: **1) Sequence tagging:** Peng-two-stage (Peng et al., 2020), OTE-MTL (Zhang et al., 2020) and JET (Xu et al., 2020). **2) Grid tagging:** GTS-BERT (Wu et al., 2020a), EMC-GCN (Chen et al., 2022a) and STAGE-3D (Liang et al., 2023). **3) MRC-based:** BMRC (Chen et al., 2021a), COM-MRC (Zhai et al., 2022) and RoBMRC (Liu et al., 2022). **4) Generative:** BARTABSA (Yan et al., 2021), LEGO-ABSA (Gao et al., 2022), EHG-Para (Lv et al., 2023) and CONTRASTE (Mukherjee et al., 2023). In addition, we investigate how well can ChatGPT (OpenAI, 2023) solve ASTE task with diverse experimental settings. More details are shown in Appendix B.

4.3 Implementation Details

We adopt the `bert-base-uncased`[2] as the sentence encoder. RGAT is stacked in 3 layers with 300 dimensions of nodes and 10 channels of edges. The hyperparameters τ, λ and γ of multi-view contrastive learning module are set to 0.05, 6 and 0.2.

[1] https://github.com/xuuuluuu/SemEval-Triplet-data

[2] https://github.com/huggingface/transformers

Category	Model	Res14			Lap14			Res15			Res16		
		P	R	F1	P	R	F1	P	R	F1	P	R	F1
Seq tagging	Peng-tow-stage♯	43.24	63.66	51.46	37.38	50.38	42.87	48.07	57.51	52.32	46.96	64.24	54.21
	OTE-MTL♮	62.00	55.97	58.71	49.53	39.22	43.42	56.37	40.94	47.13	62.88	52.10	56.96
	JET-BERT	70.56	55.94	62.40	55.39	47.33	51.04	64.45	51.96	57.53	70.42	58.37	63.83
Grid tagging	GTS-BERT♮	68.09	69.54	68.81	59.40	51.94	55.42	59.28	57.93	58.60	68.32	66.86	67.58
	EMC-GCN	71.21	72.39	71.78	61.70	56.26	58.81	61.54	62.47	61.93	65.62	71.30	68.33
	STAGE-3D	**78.58**	69.58	73.76	**71.98**	53.86	61.58	**73.63**	57.90	64.79	<u>76.67</u>	70.12	73.24
MRC-based	BMRC♮	75.61	61.77	67.99	70.55	48.98	57.82	68.51	53.40	60.02	71.20	61.08	65.75
	COM-MRC	75.46	68.91	72.01	62.35	58.16	60.17	68.35	61.24	64.53	71.55	71.59	71.57
	RoBMRC	72.51	72.73	72.62	<u>68.13</u>	57.09	62.12	65.90	<u>65.36</u>	65.63	69.98	**76.65**	73.16
Generative	ChatGPT†	47.18	53.62	50.19	34.22	47.91	39.92	42.83	57.94	49.25	43.75	57.09	49.54
	BARTABSA	65.52	64.99	65.25	61.41	56.19	58.69	61.54	62.47	61.93	65.62	71.30	68.33
	LEGO-ABSA	-	-	73.70	-	-	62.20	-	-	64.40	-	-	69.90
	EHG-Para	-	-	71.82	-	-	61.53	-	-	63.58	-	-	72.35
	CONTRASTE	73.60	<u>74.40</u>	<u>74.00</u>	64.20	<u>61.70</u>	<u>62.90</u>	65.30	**66.70**	<u>66.10</u>	72.20	<u>76.30</u>	<u>74.20</u>
Ours	**MvLFE**	<u>76.37</u>	**74.46**	**75.40**	66.12	**62.33**	**64.17**	<u>69.97</u>	64.14	**66.93**	**77.02**	73.41	**75.17**

Table 1: Experimental results on ASTE-Data-V2 dataset (Xu et al., 2020). The best results are in **bold** and the second best are <u>underlined</u>. The "♯" and "♮" mean that results are retrieved from Xu et al. (2020) and Chen et al. (2022a). The "†" denotes that ChatGPT results are obtained using 5-shot In Context Learning prompts with multi-view linguistic features by our setting. Other baseline results are derived from original papers.

For the joint training loss, the ratios β and μ are set to 0.01 and 0.1. During training, AdamW optimizer (Loshchilov and Hutter, 2017) is used with a learning rate of 2e-5 for BERT fine-tuning and 1e-3 for the other trainable parameters. The model is trained for 100 epochs with dropout rate of 0.5 and batch size of 16. For each dataset, we select the model with the best F1 scores on the development set and report the average results of five runs with different random seeds. Our model contains around 112M trainable parameters which are trained on a single NVIDIA A100-PCIE-40GB GPU with CUDA 11.0 and PyTorch 1.7.1. The average runtime for Res15 and other datasets is about 9 and 11 sec/epoch.

Task	Model	Res14	Lap14	Res15	Res16
ATE	GTS-BERT	82.76	80.44	78.27	81.13
	EMC-GCN	84.68	81.67	77.62	80.57
	STAGE-3D	84.99	82.62	81.21	83.86
	MvLFE	**86.84**	**84.71**	**83.44**	**87.46**
OTE	GTS-BERT	84.85	77.82	77.53	84.36
	EMC-GCN	85.62	78.85	78.97	85.33
	STAGE-3D	85.70	80.39	80.03	85.72
	MvLFE	**87.56**	**82.63**	**81.97**	**88.03**
AOPE	GTS-BERT	74.63	66.46	67.52	74.20
	EMC-GCN	76.33	67.94	67.26	74.15
	STAGE-3D	77.87	69.70	70.60	79.98
	MvLFE	**79.86**	**71.57**	**73.61**	**82.17**

Table 2: Test F1 scores on ATE, OTE and AOPE tasks. The baseline results are derived from Liang et al. (2023).

4.4 Main Results

The main results are reported in Table 1. Overall, our MvLFE model outperforms all baselines under the F1 metric and achieves superior precision and recall in most cases. The specific observations are that: (1) Methods based on other modeling paradigms outperform sequence tagging models, as the former considers the situation where one aspect or opinion term is associated with multiple opinion or aspect terms. (2) Compared with RoBMRC, MvLFE significantly improves F1 scores by 2.04% on average, overcoming the error propagation problem from pipeline inference in MRC-based models. (3) Our MvLFE model exceeds generative methods by an average of 1.18%~6.96% F1 scores, as the latter overlooks the reciprocity among the sentiment elements. Additionally, experiments on ChatGPT (see more details in Appendix B) reveal its limited ability to capture complex associations between multiple aspect and opinion terms. (4) Our MvLFE also modeled on grid tagging surpasses GTS-BERT and STAGE-3D by an average of 7.82% and 2.08% F1 scores, as it explores the prior indication effect of multi-view linguistic features for ASTE. Note that MvLFE outperforms EMC-GCN by large margins, while EMC-GCN also considers diverse linguistic features. We reckon MvLFE can capture the multi-hop interactions between matched aspect and opinion terms and focus on them with multi-layer RGAT but EMC-GCN only considers the one-hop association using a single layer GCN. Furthermore, different from EMC-GCN simply concatenates multi-view linguistic features, MvLFE employs multi-view contrastive learning and multi-semantic cross attention to align and aggregate them to maintain semantic consistency and capture the complementary higher-order interactions. More experimental analyses are shown in Appendix A.3.

Mode	Ablation	Res14	Lap14	Res15	Res16
Full	**MvLFE**	**75.40**	**64.17**	**66.93**	**75.17**
Module Ablation	w/o RGAT	74.16	62.86	65.49	73.81
	w/o MVCL	74.39	62.98	65.67	73.94
	w/o MSCA	74.81	63.52	66.21	74.53
RGAT Ablation	w/o DSN	75.02	63.88	66.59	74.77
	ERS-concat	75.19	63.93	66.77	74.95
	ERS-add	75.03	63.87	66.81	74.89
	w/o RC	75.13	63.92	66.71	74.96
Linguistic Features Ablation	w/o RPD	74.31	63.29	66.08	74.40
	w/o SDD	74.26	63.18	66.01	74.26
	w/o POS	73.91	62.93	65.84	73.89
	w/o SDT	73.61	62.44	65.37	73.60
	w/o SDD+RPD	73.85	62.87	65.82	74.03
	w/o POS+RPD	73.02	62.14	65.20	73.24
	w/o POS+SDD	72.87	61.82	64.93	72.89
	w/o SDT+RPD	72.04	60.97	64.09	72.10
	w/o SDT+SDD	71.83	60.70	63.75	71.91
	w/o SDT+POS	70.64	59.81	62.59	71.42
	w/o POS+SDD+RPD	70.80	59.96	62.88	71.93
	w/o SDT+SDD+RPD	70.64	59.73	62.64	71.69
	w/o SDT+POS+RPD	69.83	58.90	62.06	71.08
	w/o SDT+POS+SDD	69.71	58.83	61.62	70.50
	w/o ALL	67.47	56.38	59.61	68.01

Table 3: Ablation study (average F1 reported).

4.5 Experiments on subtasks

To further demonstrate the effectiveness of our MvLFE model, we compare it with three state-of-the-art grid tagging baselines on ATE, OTE and AOPE subtasks. The results are shown in Table 2. Note that our method can directly address these subtasks without additional modifications. Specifically, we observe that our MvLFE model achieves significant improvements over baselines on ATE and OTE task, which suggests that MvLFE can effectively explore the prior indication effect of multi-view linguistic features to extract the aspect and opinion terms. In addition, our MvLFE model also exceeds the compared models by a large margin on AOPE task. It further proves that MvLFE can capture reciprocity among the sentiment elements not only for extraction of aspect and opinion terms but also for matching the valid aspect-opinion pairs.

4.6 Ablation Study

To verify the rationality of our MvLFE, we conduct an ablation study with diverse settings and the experimental results are shown in Table 3.

Firstly, we conduct module ablation to investigate the effectiveness of different modules in our MvLFE model. Specifically, **w/o RGAT** denotes we directly use the initialized multi-view linguistic features for semantic alignment and feature fusion, which slightly decreases F1 scores by 1.34% on average, suggesting RGAT can refine multi-view linguistic features to attend matched aspect-opinion pairs. **w/o MVCL** means we remove *multi-view contrastive learning* for semantic alignment of multi-view linguistic features, resulting in an average decline of 1.17% on F1 scores, which shows MVCL can align the same word pairs represented by diverse linguistic features to maintain semantic consistency. **w/o MSCA** indicates we simply concatenate multi-view linguistic features without using *multi-semantic cross attention*. Thus, it fails to aggregate the complementary higher-order interactions between diverse linguistic features and achieves the dropping performance. Overall, each MvLFE module contributes to the entire performance on ASTE task.

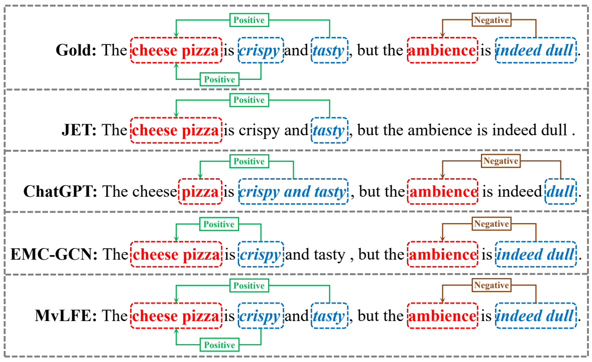

Figure 5: Case study. Aspect and opinion terms are highlighted in red and blue. The *positive* sentiment polarity is marked in green, while the *negative* in brown.

Additionally, we conduct RGAT ablation to interpret intrinsic mechanism of RGAT module. Specifically, **w/o DSN** means we utilize row normalization instead of doubly stochastic normalization (DSN) to calculate attention coefficients, which decreases F1 scores by 0.35% on average, verifying DSN can mitigate the over-smoothing issue in multi-layer RGAT learning. **ERS-concat** denotes that we use the node features h_i and h_j and the corresponding edge feature e_{ij} for concatenation as edge feature update and adding for **ERS-add**. The ablation results show the edge refinement strategy (ERS) improves by 0.21% and 0.27% on average, which indicates the ERS can consider the implicit results of aspect and opinion term extraction with enhancing the semantic connection of edge features e_{ii} and e_{jj}. **w/o RC** signifies that we remove relation constraint loss, and the degraded performance shows that it can precisely steer RGAT to refine the word-pair relations. Overall, it shows the rationality of intrinsic mechanism for RGAT module.

Lastly, we conduct linguistic features ablation to verify the role of each linguistic feature. Specifically, ablating one, two and three linguistic features results in an average performance decrease of 1.21%, 2.73% and 4.62% across all datasets, respectively. Completely removing all linguistic features (**w/o ALL**) causes a more significant performance decline by 7.55% on average. The experimental results suggest that incorporating more linguistic features leads to a more substantial improvement, likely due to the complementary role of linguistic features as the prior indication effect. Overall, it reveals that each linguistic feature is indispensable as the prior indication effect for the ASTE task can be maximized when considering all linguistic features comprehensively.

4.7 Case Study

A case study is given in Figure 5. Specifically, JET only extracts a triplet (*cheese pizza*, *tasty*, *positive*), as it cannot consider the one-to-many case. For ChatGPT, it struggles to extract aspect "*cheese pizza*" and opinion "*indeed dull*" that consist of multiple words. Moreover, it incorrectly combines two opinion terms "*crispy*" and "*tasty*" into a single opinion term. We reckon ChatGPT fails to understand the composition of sentiment elements and their association, even when given the appropriate prompts. For EMC-GCN, it ignores the triplet (*cheese pizza*, *tasty*, *positive*) as it only captures one-hop neighborhood information, while there are multi-hop connections between aspect "*cheese pizza*" and opinion "*tasty*" in syntactic dependency type. In contrast, our MvLFE model can explore the prior indication effect of multi-view linguistic features to precisely extract all sentiment triplets.

5 Conclusion

In this paper, we propose a MvLFE architecture to explore the prior indication effect of multi-view linguistic features for ASTE task in the "*Refine, Align, and Aggregate*" learning process. We first devise *relational graph attention network* to encode and refine diverse linguistic features to attend the aspect-opinion pairs. Then we employ *multi-view contrastive learning* for semantic alignment. Moreover, we utilize *multi-semantic cross attention* to aggregate them to capture the complementary higher-order interactions. Extensive experiments on benchmark datasets show the effectiveness and robustness of our MvLFE model, which consistently outperforms all baselines on ASTE task as well as several subtasks. In the future, we plan to extend our model for information extraction.

Limitations

Despite obtaining state-of-the-art performance, our proposed approach still has some following limitations to consider for potential future directions.

- Our MvLFE model is constructed using the grid tagging paradigm. Although grid tagging methods can effectively capture the association among sentiment elements, they commonly suffer from a limitation, i.e., most of the relation labels between two words are irrelevant to sentiment elements (an example shown in Figure 3), which leads to the class imbalance problem of word-pair relations. Thus, when most irrelevant labels are predicted correctly but few for sentiment-related labels, it causes the phenomenon that the overall loss is low but limited F1 scores of sentiment triplet extraction, which potentially misguides the parameter optimization. Additionally, the requirement to construct a word-pair relations table takes up more memory compared to the models that only consider sequence representations.

- Experiments only verified the consistent improvement on ASTE task, while we intuitively reckon the idea of MvLFE that exploits the prior indication effect of multi-view linguistic features in the "*Refine, Align, and Aggregate*" learning process can be expanded to some triplet extraction tasks, such as *Relation Extraction* and *Event Extraction*.

- The performance of our MvLFE model is still somewhat affected by the parsing quality of the CoreNLP toolkit. The good news is that the CoreNLP toolkit has demonstrated its effectiveness in syntactic dependency analysis and part-of-speech tagging. Thus, despite this limitation, we can still utilize the CoreNLP toolkit to support our research.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62272188) and the Fundamental Research Funds for the Central Universities 2662021JC008. We appreciate the insightful comments and constructive feedback from the anonymous reviewers. We also thank the Huazhong Agricultural University for providing the experimental resources.

References

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Deli Chen, Yankai Lin, Wei Li, Peng Li, Jie Zhou, and Xu Sun. 2020a. Measuring and relieving the over-smoothing problem for graph neural networks from the topological view. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 3438–3445. AAAI Press.

Hao Chen, Zepeng Zhai, Fangxiang Feng, Ruifan Li, and Xiaojie Wang. 2022a. Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2974–2985. Association for Computational Linguistics.

Shaowei Chen, Jie Liu, Yu Wang, Wenzheng Zhang, and Ziming Chi. 2020b. Synchronous double-channel recurrent network for aspect-opinion pair extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 6515–6524. Association for Computational Linguistics.

Shaowei Chen, Yu Wang, Jie Liu, and Yuelin Wang. 2021a. Bidirectional machine reading comprehension for aspect sentiment triplet extraction. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 12666–12674. AAAI Press.

Yuqi Chen, Keming Chen, Xian Sun, and Zequn Zhang. 2022b. A span-level bidirectional network for aspect sentiment triplet extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 4300–4309. Association for Computational Linguistics.

Zhexue Chen, Hong Huang, Bang Liu, Xuanhua Shi, and Hai Jin. 2021b. Semantic and syntactic enhanced aspect sentiment triplet extraction. In *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021, Online Event, August 1-6, 2021*, volume ACL/IJCNLP 2021 of *Findings of ACL*, pages 1474–1483. Association for Computational Linguistics.

Zhuang Chen and Tieyun Qian. 2020. Relation-aware collaborative learning for unified aspect-based sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 3685–3694. Association for Computational Linguistics.

Hongliang Dai and Yangqiu Song. 2019. Neural aspect and opinion term extraction with mined rules as weak supervision. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 5268–5277. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 4171–4186. Association for Computational Linguistics.

Haiwen Diao, Ying Zhang, Lin Ma, and Huchuan Lu. 2021. Similarity reasoning and filtration for image-text matching. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 1218–1226. AAAI Press.

Timothy Dozat and Christopher D. Manning. 2017. Deep biaffine attention for neural dependency parsing. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net.

Fartash Faghri, David J. Fleet, Jamie Ryan Kiros, and Sanja Fidler. 2017. VSE++: improved visual-semantic embeddings. *CoRR*, abs/1707.05612.

Zhifang Fan, Zhen Wu, Xin-Yu Dai, Shujian Huang, and Jiajun Chen. 2019. Target-oriented opinion words extraction with target-fused neural sequence labeling. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, pages 2509–2518. Association for Computational Linguistics.

Hao Fei, Fei Li, Chenliang Li, Shengqiong Wu, Jingye Li, and Donghong Ji. 2022. Inheriting the wisdom of predecessors: A multiplex cascade framework for unified aspect-based sentiment analysis. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI*, pages 4096–4103.

Lei Gao, Yulong Wang, Tongcun Liu, Jingyu Wang, Lei Zhang, and Jianxin Liao. 2021. Question-driven span labeling model for aspect-opinion pair extraction. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 12875–12883. AAAI Press.

Tianhao Gao, Jun Fang, Hanyu Liu, Zhiyuan Liu, Chao Liu, Pengzhang Liu, Yongjun Bao, and Weipeng Yan. 2022. LEGO-ABSA: A prompt-based task assemblable unified generative framework for multi-task aspect-based sentiment analysis. In *Proceedings of the 29th International Conference on Computational Linguistics, COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*, pages 7002–7012. International Committee on Computational Linguistics.

Zhibin Gou, Qingyan Guo, and Yujiu Yang. 2023. Mvp: Multi-view prompting improves aspect sentiment tuple prediction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 4380–4397. Association for Computational Linguistics.

Ruidan He, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2019. An interactive multi-task learning network for end-to-end aspect-based sentiment analysis. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 504–515. Association for Computational Linguistics.

Hongjiang Jing, Zuchao Li, Hai Zhao, and Shu Jiang. 2021. Seeking common but distinguishing difference, A joint aspect-based sentiment analysis model. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 3910–3922. Association for Computational Linguistics.

Kuang-Huei Lee, Xi Chen, Gang Hua, Houdong Hu, and Xiaodong He. 2018. Stacked cross attention for image-text matching. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part IV*, volume 11208 of *Lecture Notes in Computer Science*, pages 212–228. Springer.

Kun Li, Chengbo Chen, Xiaojun Quan, Qing Ling, and Yan Song. 2020. Conditional augmentation for aspect term extraction via masked sequence-to-sequence generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 7056–7066. Association for Computational Linguistics.

Shuo Liang, Wei Wei, Xian-Ling Mao, Yuanyuan Fu, Rui Fang, and Dangyang Chen. 2023. STAGE: span tagging and greedy inference scheme for aspect sentiment triplet extraction. In *Thirty-Seventh AAAI Conference on Artificial Intelligence, AAAI 2023, Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence, IAAI 2023, Thirteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2023, Washington, DC, USA, February 7-14, 2023*, pages 13174–13182. AAAI Press.

Shu Liu, Kaiwen Li, and Zuhe Li. 2022. A robustly optimized BMRC for aspect sentiment triplet extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL 2022, Seattle, WA, United States, July 10-15, 2022*, pages 272–278. Association for Computational Linguistics.

Ilya Loshchilov and Frank Hutter. 2017. Fixing weight decay regularization in adam. *CoRR*, abs/1711.05101.

Haoran Lv, Junyi Liu, Henan Wang, Yaoming Wang, Jixiang Luo, and Yaxiao Liu. 2023. Efficient hybrid generation framework for aspect-based sentiment analysis. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2023, Dubrovnik, Croatia, May 2-6, 2023*, pages 1007–1018. Association for Computational Linguistics.

Dehong Ma, Sujian Li, Fangzhao Wu, Xing Xie, and Houfeng Wang. 2019. Exploring sequence-to-sequence learning in aspect term extraction. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, pages 3538–3547. Association for Computational Linguistics.

Fukun Ma, Xuming Hu, Aiwei Liu, Yawen Yang, Shuang Li, Philip S. Yu, and Lijie Wen. 2023. Amr-based network for aspect-based sentiment analysis. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 322–337. Association for Computational Linguistics.

Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. 2014. The stanford corenlp natural language processing toolkit. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014, June 22-27, 2014, Baltimore, MD, USA, System Demonstrations*, pages 55–60. The Association for Computer Linguistics.

Yue Mao, Yi Shen, Jingchao Yang, Xiaoying Zhu, and Longjun Cai. 2022. Seq2path: Generating sentiment tuples as paths of a tree. In *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2215–2225. Association for Computational Linguistics.

Yue Mao, Yi Shen, Chao Yu, and Longjun Cai. 2021. A joint training dual-mrc framework for aspect based sentiment analysis. In *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*, pages 13543–13551. AAAI Press.

Rajdeep Mukherjee, Nithish Kannen, Saurabh Kumar Pandey, and Pawan Goyal. 2023. CONTRASTE: supervised contrastive pre-training with aspect-based prompts for aspect sentiment triplet extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2023, Singapore, December 6-10, 2023*, pages 12065–12080. Association for Computational Linguistics.

OpenAI. 2023. Introducing chatgpt.

Zhengxin Pan, Fangyu Wu, and Bailing Zhang. 2023. Fine-grained image-text matching by cross-modal hard aligning network. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023*, pages 19275–19284. IEEE.

Haiyun Peng, Lu Xu, Lidong Bing, Fei Huang, Wei Lu, and Luo Si. 2020. Knowing what, how and why: A near complete solution for aspect-based sentiment analysis. In *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 8600–8607. AAAI Press.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammad Al-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, Véronique Hoste, Marianna Apidianaki, Xavier Tannier, Natalia V. Loukachevitch, Evgeniy V. Kotelnikov, Núria Bel, Salud María Jiménez Zafra, and Gülsen Eryigit. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2016, San Diego, CA, USA, June 16-17, 2016*, pages 19–30. The Association for Computer Linguistics.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. Semeval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2015, Denver, Colorado, USA, June 4-5, 2015*, pages 486–495. The Association for Computer Linguistics.

Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.

Xipeng Qiu, Tianxiang Sun, Yige Xu, Yunfan Shao, Ning Dai, and Xuanjing Huang. 2020. Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, 63(10):1872–1897.

Siyu Tang, Heyan Chai, Ziyi Yao, Ye Ding, Cuiyun Gao, Binxing Fang, and Qing Liao. 2022. Affective knowledge enhanced multiple-graph fusion networks for aspect-based sentiment analysis. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 5352–5362. Association for Computational Linguistics.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 5998–6008.

Petar Velickovic, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò, and Yoshua Bengio. 2018. Graph attention networks. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net.

Bo Wang, Armin Pourshafeie, Marinka Zitnik, Junjie Zhu, Carlos D Bustamante, Serafim Batzoglou, and Jure Leskovec. 2018. Network enhancement as a general method to denoise weighted biological networks. *Nature communications*, 9(1):3108.

Qianlong Wang, Zhiyuan Wen, Qin Zhao, Min Yang, and Ruifeng Xu. 2021. Progressive self-training with discriminator for aspect term extraction. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021*, pages 257–268. Association for Computational Linguistics.

Zhen Wu, Chengcan Ying, Fei Zhao, Zhifang Fan, Xinyu Dai, and Rui Xia. 2020a. Grid tagging scheme for aspect-oriented fine-grained opinion extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2576–2585.

Zhen Wu, Fei Zhao, Xin-Yu Dai, Shujian Huang, and Jiajun Chen. 2020b. Latent opinions transfer network for target-oriented opinion words extraction. In *The*

Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*, pages 9298–9305. AAAI Press.

Lu Xu, Yew Ken Chia, and Lidong Bing. 2021. Learning span-level interactions for aspect sentiment triplet extraction. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 4755–4766. Association for Computational Linguistics.

Lu Xu, Hao Li, Wei Lu, and Lidong Bing. 2020. Position-aware tagging for aspect sentiment triplet extraction. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*, pages 2339–2349. Association for Computational Linguistics.

Hang Yan, Junqi Dai, Tuo Ji, Xipeng Qiu, and Zheng Zhang. 2021. A unified generative framework for aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1-6, 2021*, pages 2416–2429. Association for Computational Linguistics.

Guoxin Yu, Lemao Liu, Haiyun Jiang, Shuming Shi, and Xiang Ao. 2023. Making better use of training corpus: Retrieval-based aspect sentiment triplet extraction via label interpolation. In *Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 4914–4927. Association for Computational Linguistics.

Zepeng Zhai, Hao Chen, Fangxiang Feng, Ruifan Li, and Xiaojie Wang. 2022. COM-MRC: A context-masked machine reading comprehension framework for aspect sentiment triplet extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 3230–3241. Association for Computational Linguistics.

Chen Zhang, Qiuchi Li, Dawei Song, and Benyou Wang. 2020. A multi-task learning framework for opinion triplet extraction. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*, volume EMNLP 2020 of *Findings of ACL*, pages 819–828. Association for Computational Linguistics.

Mao Zhang, Yongxin Zhu, Zhen Liu, Zhimin Bao, Yunfei Wu, Xing Sun, and Linli Xu. 2023. Span-level aspect-based sentiment analysis via table filling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 9273–9284. Association for Computational Linguistics.

Wenxuan Zhang, Xin Li, Yang Deng, Lidong Bing, and Wai Lam. 2021. Towards generative aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 2: Short Papers), Virtual Event, August 1-6, 2021*, pages 504–510. Association for Computational Linguistics.

Yice Zhang, Yifan Yang, Yihui Li, Bin Liang, Shiwei Chen, Yixue Dang, Min Yang, and Ruifeng Xu. 2022. Boundary-driven table-filling for aspect sentiment triplet extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 6485–6498. Association for Computational Linguistics.

He Zhao, Longtao Huang, Rong Zhang, Quan Lu, and Hui Xue. 2020. Spanmlt: A span-based multi-task learning framework for pair-wise aspect and opinion terms extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*, pages 3239–3248. Association for Computational Linguistics.

Yichun Zhao, Kui Meng, Gongshen Liu, Jintao Du, and Huijia Zhu. 2022. A multi-task dual-tree network for aspect sentiment triplet extraction. In *Proceedings of the 29th International Conference on Computational Linguistics, COLING 2022, Gyeongju, Republic of Korea, October 12-17, 2022*, pages 7065–7074. International Committee on Computational Linguistics.

A Additional Details

A.1 Triplets Decoding

The details of sentiment triplets decoding are shown in Algorithm 1. Specifically, we first obtain the set of diagonal tags $\mathcal{D}$ from the prediction results. Each tag in $\mathcal{D}$ represents the relation between a word and itself. Then we extract each aspect term or opinion term based on consecutive occurrences of "B-A" and "I-A" labels or "B-O" and "I-O" labels in the diagonal tags set $\mathcal{D}$, thereby forming the aspects set $\mathcal{A}$ and the opinions set $\mathcal{O}$. Finally, we identify the sentiment polarities between the matched aspect and opinion terms based on their predicted relations, thus yielding the sentiment triplets set $\mathcal{T}$.

Algorithm 1 Triplet Decoding for ASTE

Input: The predictions $\mathcal{P} = \{p_{11}, p_{12}, \cdots, p_{nn}\}$ of the sentence X with n words. p_{ij} denotes the tag label of the word pair (w_i,w_j).

Output: Triplets set $\mathcal{T}$ of the given sentence.

Initialize aspects set $\mathcal{A} = \{\}$, opinions set $\mathcal{O} = \{\}$, diagonal tags of predictions $\mathcal{D} = \{\}$ and triplets set $\mathcal{T} = \{\}$.
 # Get diagonal tags of predictions
while $i \leq n$ **do**
    $\mathcal{D}$.append(p_{ii}), $i \leftarrow i + 1$
end while
 # Get aspects
for p_{ii} in $\mathcal{D}$ **do**
    if p_{ii} =="B-A" **then**
        while $(i + 1) \leq n$ **do**
            if $p_{i+1,i+1}$! ="I-A" **then** $j \leftarrow i$
            end if
        end while
        $\mathcal{A}$.append($[w_i, w_{i+1}, \cdots, w_j]$)
    end if
end for
 # Get opinions
for p_{ii} in $\mathcal{D}$ **do**
    if p_{ii} =="B-O" **then**
        while $(i + 1) \leq n$ **do**
            if $p_{i+1,i+1}$! ="I-O" **then** $j \leftarrow i$
            end if
        end while
        $\mathcal{O}$.append($[w_i, w_{i+1}, \cdots, w_j]$)
    end if
end for
 # Get sentiment triplets
while $a \in \mathcal{A}$ and $o \in \mathcal{O}$ **do**
    $\mathcal{S} = \{\}$
    while $w_i \in a$ and $w_j \in o$ **do**
        if $i < j$ **then** $tag = p_{ij}$
        else $tag = p_{ij}$
        end if
        if $tag \in \{POS, NEU, NEG\}$ **then**
            $\mathcal{S} \leftarrow \mathcal{S} \cup (tag)$
        end if
    end while
    if $\mathcal{S} \neq \{\}$ **then**
        $s = argmax(\mathcal{S}), \mathcal{T} \leftarrow \mathcal{T} \cup \{(a, o, s)\}$
    end if
end while

A.2 Dataset Statistics

Table 4 shows details about the experiment dataset. We have counted the number of sentences, sentiment triplets for each sentiment polarity, aspect terms and opinion terms within each dataset.

A.3 Additional Experiments and Analyses

A.3.1 Visualization of Linguistic Features

To intuitively reflect the prior indication effect for the ASTE task, we further visualize word-pair relations represented by multi-view linguistic features after refining with different RGAT layers, by utilizing l_2 norm to normalize the edge features into relevance coefficients of the corresponding word pairs.

As shown in Figure 6, the sampled sentence "*Tasty food but poor service*" contains two sentiment triplets (*food*, *tasty*, *positive*) and (*service*, *poor*, *negative*), we can conclude by specific ob-

Dataset		#S	#T	#POS	#NEU	#NEG	#A	#O
Res14	Train	1266	2337	1691	166	480	2051	2061
	Dev	310	577	404	54	119	500	497
	Test	492	994	773	66	155	844	994
Lap14	Train	906	1460	817	126	517	1254	1460
	Dev	219	345	169	36	140	302	346
	Test	328	541	364	63	114	466	543
Res15	Train	605	1013	783	25	205	935	1013
	Dev	148	249	185	11	53	236	249
	Test	322	485	317	25	143	460	485
Res16	Train	857	1394	1015	50	329	1300	1394
	Dev	210	339	252	11	76	319	339
	Test	326	514	407	29	78	474	514

Table 4: Statistics for ASTE-Data-V2 dataset. #S and #T mean the total number of sentences and triplets. #POS, #NEU and #NEG denote the number of positive, neutral, and negative sentiment triplets respectively. #A and #O represent the number of aspect and opinion terms.

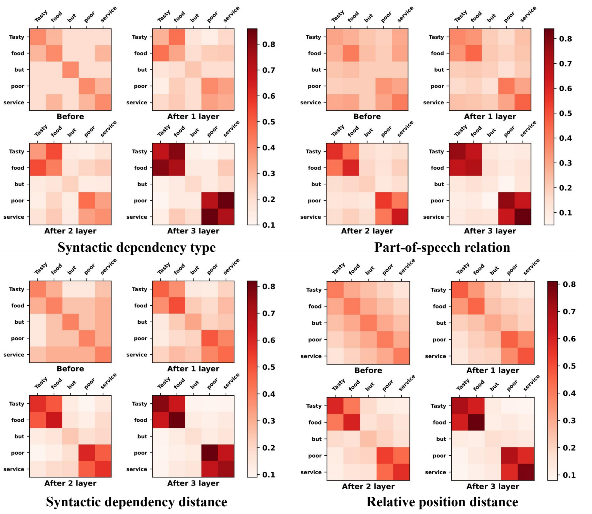

Figure 6: Visualization of multi-view linguistic features after refining with different RGAT layers.

servations from the visualization: (1) From the visualization of syntactic dependency type, our MvLFE model pays more attention to word-pair relations with "*nsubj*" syntactic type across multi-layer RGAT refinement, because aspect terms are typically the nominal subject of opinion terms. (2) From the visualization of part-of-speech relation, the correlation coefficients between nouns and adjectives are higher when RGAT deepens, as nouns and adjectives are more likely to form matching aspect-opinion pairs. (3) From the visualization of syntactic dependency distance and relative position distance, the connections between matched aspect-opinion pairs are tighter after multi-layer RGAT learning, as valid aspect-opinion pairs are closer in syntactic and relative distance.

In summary, each view of linguistic feature contributes to the unique prior indication effect for the ASTE task, which is indispensable and effective.

A.3.2 Hyperparameters Analysis

To investigate the impact of some major hyperparameters, we conduct sensitivity analysis as illustrated in Figure 7.

Specifically, our MvLFE model achieves optimal performance when RGAT is three layers with a range from single to five layers, where multiple layers indicate that MvLFE can capture multi-hop neighborhood information between words. We reckon that 3-hop interaction between words helps to match the valid aspect and opinion terms with vast distance. Note that the performance declines as the RGAT goes deeper, which may be due to the model matching the invalid aspects and opinions.

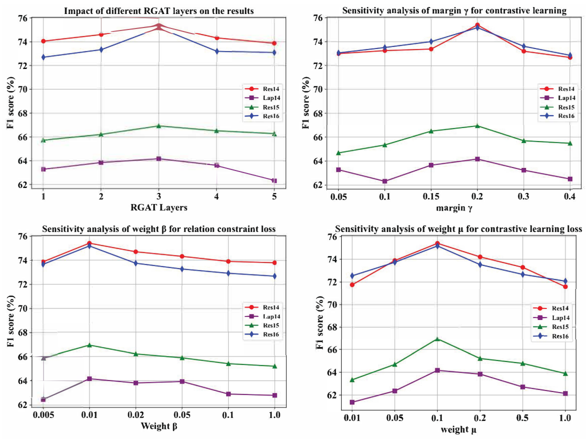

Figure 7: Sensitivity analysis of hyperparameters.

Model	PLM	Res14	Lap14	Res15	Res16
MvLFE	BERT	75.40	64.17	66.93	75.17
	RoBERTa	75.36	64.31	66.87	75.42
	XLNet	75.61	64.32	66.82	75.49
	ALBERT	75.19	64.21	66.79	75.24

Table 5: Comparison results with different backbones.

For other hyperparameters, margin γ is a threshold that ensures the minimum semantic distance between positive and negative sample pairs to properly align linguistic features with semantic features. β and μ are weights used to adjust relation constraint loss and contrastive learning loss to measure the impact on the training objective. As they gradually increase, the performance initially improves and then decreases, and we finally set γ, β and μ to 0.2, 0.01 and 0.1 as optimal selection.

A.3.3 Experiments on Different Backbones

We conduct additional experiments to analyze the sensitivity of our MvLFE model using different pre-trained language models as sentence encoders. The experimental results in Table 5 demonstrate that the performance of our model does not exhibit significant variations when employing different pre-trained language models, which suggests the robustness of our model. To ensure fairness, we adopt BERT as the sentence encoder the same as most of the related studies.

A.3.4 Potential Practical Applications

Time complexity: Our MvLFE model is quadratic relative to the input data. The primary source of complexity in this quadratic time complexity is the attention operations within the transformer.

Space complexity: Our MvLFE model takes up an additional parameter space occupation amounts to

Model	Res14			Lap14			Res15			Res16		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
zero-shot w/o task instruction	13.96	18.42	15.88	11.28	13.59	12.33	12.36	14.84	13.49	14.37	16.33	15.29
zero-shot w/ task instruction	21.07	24.16	22.51	18.74	21.92	20.21	20.78	23.69	22.14	22.65	24.94	23.74
1-shot ICL	25.54	29.37	27.32	21.29	24.79	22.91	24.91	28.85	26.74	26.23	30.08	28.02
1-shot ICL w/ MvLF	31.24	33.98	32.55	26.38	29.49	27.85	30.72	31.92	31.31	32.47	33.81	33.13
3-shot ICL	30.82	33.71	32.20	27.01	29.43	28.17	30.69	31.97	31.32	31.45	33.01	32.21
3-shot ICL w/ MvLF	40.72	45.63	43.04	31.09	37.24	33.89	37.81	45.96	41.49	39.23	47.81	43.10
5-shot ICL	41.27	46.09	43.55	31.16	37.33	33.97	37.98	46.83	41.94	39.55	48.49	43.57
5-shot ICL w/ MvLF	**47.18**	**53.62**	**50.19**	**34.22**	**47.91**	**39.92**	**42.83**	**57.94**	**49.25**	**43.75**	**57.09**	**49.54**

Table 6: Comparing ASTE results obtained using ChatGPT with different prompt settings. ICL means in-context learning and MvLF denotes multi-view linguistic features.

2M derived from constructing the word-pair relations table, which is notably minor when compared to the parameter size of the BERT model (110M).

B Experiments with ChatGPT

Recently, Large Language Models (LLMs) such as ChatGPT have sparked a revolutionary change in natural language processing technology. They are capable of achieving impressive in-context learning (ICL) (Brown et al., 2020) results with zero-shot and few-shot prompts for unseen tasks, without the need for any parameter updates.

In this paper, we carried out some experiments on four benchmark datasets to investigate how well can ChatGPT solve ASTE task with diverse prompt settings. The experimental results are shown in Table 6. Moreover, we list some examples of prompts for the experimental settings in Table 7. Note that we only list the zero-shot and one-shot prompts due to the limited length of the table. The only difference between few-shot prompts and one-shot prompts is the presence of more data samples.

Specifically, we employ the following experimental settings to verify the performance of ChatGPT on the ASTE task: (1) zero-shot without task instruction, (2) zero-shot with task instruction, (3) one-shot ICL, (4) one-shot ICL with multi-view linguistic features, (5) 3-shot ICL, (6) 3-shot ICL with multi-view linguistic features, (7) 5-shot ICL and (8) 5-shot ICL with multi-view linguistic features. From the experimental results, we first observe that using the task instruction significantly improves the average F1 scores by 7.90%, which suggests it is effective in enabling ChatGPT to understand the purpose of the ASTE task. Second, prompting with training samples further improves the performance. We suppose the reason is that ChatGPT is able to learn from more training samples to extract sentiment triplets from different situations. Note that we also conduct additional experiments for each ICL setting by incorporating multi-view linguistic features of the given samples. The experimental results show that ChatGPT improves by a large margin with the addition of multi-view linguistic features, suggesting the effectiveness of the prior indication effect of multi-view linguistic features for ASTE.

While ChatGPT has achieved promising results, there is still a gap between ChatGPT and SOTA methods on ASTE task. Based on our observation of the output results, ChatGPT has several limitations to solving ASTE task: (1) It tends to extract incomplete aspect and opinion terms. (2) It often merges multiple juxtaposed opinion terms into a single opinion term. (3) It is prone to under-extracting in situations where multiple sentiment triplets exist.

Setting	Prompt
zero-shot w/o task instruction	**Given the review:** extract all the sentiment triplets in the review and return the result in JSON format. Remember that no explanation is required and there should be no irrelevant text replies! **Review:** The cheese pizza is crispy and tasty, but the ambience is indeed dull.
zero-shot w/ task instruction	**Given the review for ASTE task:** please follow the task instruction and return the result in JSON format. Remember that no explanation is required and there should be no irrelevant text replies! **Task instruction:** ASTE aims to extract all sentiment triplets from the review, and each triplet contains three elements, namely aspect term, opinion term and their associated sentiment. The sentiment polarity belongs to the set {positive, neutral, negative}. **Review:** The cheese pizza is crispy and tasty, but the ambience is indeed dull.
one-shot ICL	**Given the review and several extraction samples for ASTE task:** please follow the task instruction and several extraction samples, then return the result in JSON format. Remember that no explanation is required and there should be no irrelevant text replies! **Task instruction:** ASTE aims to extract all sentiment triplets from the review, and each triplet contains three elements, namely aspect term, opinion term and their associated sentiment. The sentiment polarity belongs to the set {positive, neutral, negative}. **Sample1:** The food was just OK, at least for what food was available. **Result of sample1:** Sentiment triplets: (food, ok, positive) **Review:** The cheese pizza is crispy and tasty, but the ambience is indeed dull.
one-shot ICL w/ MvLF	**Given the review and several extraction samples with multi-view linguistic features for ASTE task:** please follow the task instruction and several extraction samples with multi-view linguistic features, then return the result in JSON format. Remember that no explanation is required and there should be no irrelevant text replies! **Task instruction:** ASTE aims to extract all sentiment triplets from the review, and each triplet contains three elements, namely aspect term, opinion term and their associated sentiment. The sentiment polarity belongs to the set {positive, neutral, negative}. **The prior indication effect of multi-view linguistic features for ASTE:** (1) in the view of syntactic dependency type, the aspect term is the nominal subject ("nsubj") of the opinion term, also the opinion term is adjectival modifier ("amod") of the the aspect term. (2) in the view of part-of-speech relation, aspect terms are nouns while opinion terms are adjectives. Hence, the word pair with the "NN-JJ" part-of-speech combination tends to form an aspect-opinion pair. (3) in the view of syntactic dependency distance, the matched aspect and opinion terms are closer. (4) in the view of relative position distance, the matched aspect and opinion terms are closer. **Sample1:** The food was just OK, at least for what food was available. **Multi-view linguistic features for sample1:** We provide part-of-speech tagging labels for each word in the review, as well as syntactic dependency type between each word and its parent word. The "Head" indicates the position of the parent word in the review (starting from 1, with 0 representing the root node). Based on this information, you need to calculate the syntactic dependency type, part-of-speech relation, syntactic dependency distance, and relative position distance between words to obtain the multi-view linguistic features of the given comment. **Part-of-speech tagging labels:** ["DT", "NN", "VBD", "RB", "JJ", ",", "RB", "RBS", "IN", "WDT", "NN", "VBD", "JJ", "."] **Syntactic dependency type:** ["det", "nsubj", "cop", "advmod", "root", "punct", "advmod", "fixed", "mark", "det", "nsubj", "cop", "advcl", "punct"] **Head:** [2, 5, 5, 5, 0, 5, 13, 7, 13, 11, 13, 13, 5, 5] **Result of sample1:** Sentiment triplets: (food, ok, positive) **Review:** The cheese pizza is crispy and tasty, but the ambience is indeed dull.

Table 7: Some prompts for ChatGPT.