RESEARCH ARTICLE

Aspect Sentiment Triplet Extraction with Syntax-Semantics Graph Convolutional Network

Jingyun Zhang[1,2] · Shuwei Xu[1,2] · Xin Gao[1,2] · Zhiwei Tang[1,2]

Received: 29 December 2024 / Revised: 1 June 2025 / Accepted: 17 June 2025
© The Author(s) 2025

Abstract
In the traditional task of aspect sentiment triplet extraction, existing approaches typically focus on either syntactic or semantic features independently, failing to leverage the complementary integration of these two types of information. Although graph convolutional network-based approaches have demonstrated impressive performance in triplet extraction tasks, they often ignore distance features and semantic information when capturing sentence information. As a result, the integration of syntactic and semantic information remains suboptimal, negatively impacting sentiment analysis performance. To address this limitation, we propose a novel Syntax-Semantics Graph Convolutional Network for aspect sentiment triplet extraction. Our method first extracts syntactic structural information using the probability matrix of dependency trees, from which a mask matrix is constructed based on the varying distances between words. Next, semantic information is captured via a self-attention mechanism and an aspect-attention mechanism, utilizing an attention score matrix. Finally, an interaction module is introduced to effectively integrate syntactic and semantic features. Extensive experiments on several benchmark datasets demonstrate that our approach significantly outperforms existing baselines, achieving an average F1-score improvement of at least 1.083%.

Keywords Aspect-based sentiment analysis · Aspect sentiment triplet extraction · Graph convolutional network · Syntactic information and semantic information

1 Introduction

With the proliferation of social media, online reviews, and question-and-answer platforms, user-generated content has become increasingly prevalent. Such content is often concise and may discuss multiple targets within a single sentence, posing challenges for traditional sentiment analysis methods. These methods typically provide an overall sentiment for the sentence but fail to capture sentiments toward specific aspects. To address this limitation, aspect-based sentiment analysis (ABSA) has emerged, with one of its core tasks being aspect sentiment triplet extraction (ASTE). ASTE involves identifying three key elements within a sentence: aspect terms, opinion terms, and their associated sentiment polarity [1]. Specifically, aspect terms refer to the subjects or targets discussed or evaluated. Opinion terms are words expressing sentiments (e.g., "good," "bad," "satisfied," "dissatisfied"). Sentiment polarity denotes the direction of the sentiment, typically categorized as positive, negative, or neutral. Figure 1 provides a visual illustration of the triplet extraction process.

Aspect sentiment triplet extraction, a complex task in aspect-based sentiment analysis, presents several challenges, such as polysemy ambiguity and long-distance dependencies. To tackle these challenges, researchers have incorporated syntactic and semantic information. Syntactic information helps uncover relationships between aspect terms and opinion terms by analyzing sentence structure and dependency relations (e.g., subject–predicate–object).



Published online: 01 July 2025

 https://doi.org/10.1007/s44196-025-00900-w

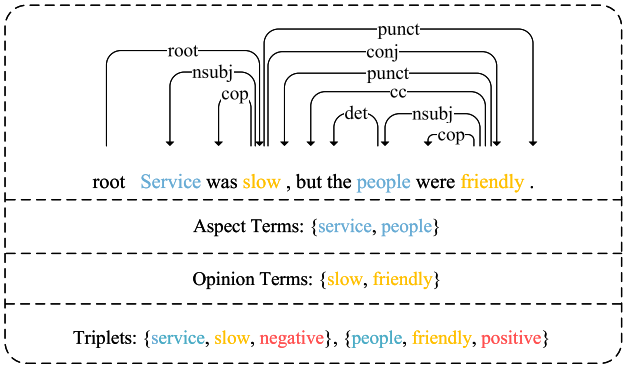


Fig. 1 ASTE task (aspect terms, opinion terms, and sentiment polarity highlighted in blue, green, and red, respectively)

It enables identifying word dependencies and parsing complex sentence structures, which is critical for accurately extracting multiple aspect terms and their sentiment polarities. Semantic information, in contrast, aids in resolving polysemy ambiguity and identifying implicit aspects or sentiments by capturing word meanings and contextual nuances. This is especially important for determining the correct sentiment polarity in ambiguous or context-dependent cases. Studies have demonstrated that integrating syntactic and semantic features can significantly enhance performance in sentiment analysis tasks [2].

To achieve a comprehensive integration of syntactic and semantic information, this paper proposes a Syntax-Semantics Graph Convolutional Network (SSGCN) for aspect sentiment triplet extraction. First, syntactic information extraction: A dependency tree probability matrix is employed to capture rich syntactic features, complemented by a mask matrix constructed based on word distances. The dependency tree effectively represents the syntactic structure of a sentence, while the probability matrix ensures comprehensive coverage of syntactic structure information. Then, semantic information extraction: An attention score matrix, combining the self-attention and aspect-attention mechanisms, is utilized to incorporate semantic features. The self-attention mechanism captures global contextual dependencies, while the aspect-attention mechanism focuses on specific aspect terms, enhancing the understanding of sentiment contexts. Finally, an interaction module is constructed to fully integrate syntactic information and semantic information. We have conducted extensive experiments on public datasets, and the experimental results show that SSGCN can fully integrate syntactic information and semantic information and has better performance than other baseline methods.

2 Related Work

Aspect-based sentiment analysis is a fine-grained sentiment analysis task that focuses on identifying the sentiment polarity of specific aspects within a sentence. ABSA encompasses three fundamental subtasks, Aspect Term Extraction (ATE) [3–6], Opinion Term Extraction (OTE) [7–9], and Aspect Sentiment Classification (ASC) [10–14]. Traditional approaches typically process these subtasks independently, overlooking their inherent interdependencies. To address this limitation, recent studies have shifted toward joint modeling of multiple subtasks to explore their interrelationships. This shift has led to the emergence of several composite tasks. Aspect–Opinion Pair Extraction (AOPE) [15–17] identifies aspect terms in a sentence and their associated opinion terms. Aspect Sentiment Pair Extraction (ASPE) [18–20] identifies aspect terms and their corresponding sentiment polarity pairs. Aspect Category Sentiment Analysis (ACSA) [21, 22] extends the scope to detect aspect categories discussed and their sentiment polarities. Despite these advancements, most existing studies fail to simultaneously extract complete information, including aspect terms, opinion terms, and sentiment polarity, within a unified framework. To overcome this challenge, Peng et al. [1] introduced the Aspect Sentiment Triplet Extraction (ASTE) task. They proposed a two-stage pipeline method. In the first stage, the method extracted aspect terms and opinion terms with sentiment attributes, respectively, through two sequence labeling models. In the second stage, a classifier was used to identify valid and paired aspect–opinion terms from the predicted aspect terms and opinion terms, and finally, triple predictions was constructed.

In recent years, researchers have increasingly leveraged syntactic information, particularly dependency parsing, to enhance ASTE task performance by capturing structural relationships in sentences [23–25]. Wang et al. employed dependency parsing and syntactic tree information to construct domain-adaptive syntactic representations, improving cross-domain robustness [26]. Xu et al. proposed a position-aware tagging scheme that, while focused on position encoding, highlighted the auxiliary role of syntactic information in contextual understanding [27]. Wu et al. introduced a grid tagging scheme that integrated dependency trees to address complex relationships, enhancing the extraction of aspect–opinion term connections [28]. With the rise of Transformer-based models, researchers have also begun exploring methods to combine syntactic information with Transformer architectures. For example, Yuan et al. innovatively incorporated dependency parsing graphs and graph neural networks (GNNs) into the Transformer model. This hybrid design enabled the model to encode syntactic information more effectively, leading to improved performance in ASTE tasks [29].

Semantic information plays a critical role in the ASTE task, with many studies utilizing attention mechanisms to capture it effectively [30–32]. For instance, Jiang et al. proposed a semantically enhanced dual-encoder architecture, where attention mechanisms were employed to process the contextual information of aspect terms and sentiment, strengthening the semantic relationship between sentence elements and improving triplet extraction accuracy [33]. Similarly, Chen et al. introduced a multi-task learning framework that combined shared layers for global feature extraction and task-specific layers for handling individual subtasks. The shared layers primarily captured semantic information, which supported subsequent task processing [34]. To further enhance semantic understanding, Chen et al. incorporated a bidirectional machine reading comprehension framework that combined a question–answering mechanism with deep semantic representations, improving the handling of complex sentences [35]. Additionally, an enhanced multi-channel graph convolutional network model was proposed to fully utilize word relationships, further advancing semantic feature extraction [36]. Despite these advancements, current models face challenges in handling polysemy and complex contexts. These limitations stem largely from the separate processing of syntactic and semantic information, which prevents effective integration. To address this, we propose a unified framework that comprehensively integrates syntactic and semantic information from word vectors, enabling complementary interactions and overcoming the shortcomings of previous approaches.

3 Model

Given a sentence $X = \{w_1, w_2, \ldots, w_n\}$ containing n words, the task is to extract a set of triples $T = \{(a, o, s)_m\}_{m=1}^{|T|}$, where a denotes aspect term, o denotes opinion term, and s is the sentiment polarity of the aspect term. The sentiment polarity of aspect term belongs to the set $S = \{\text{POS}, \text{NEU}, \text{NEG}\}$, representing positive, neutral, and negative sentiments, respectively. Sentence X contains $|T|$ triples. The overall framework of the model is shown in Fig. 2.

3.1 Input and Encoding Layer

We use BERT as a sentence encoder to obtain the hidden context representation of each word in the sentence [37]. Given a sentence with n words $X = \{w_1, w_2, \ldots, w_n\}$, the BERT layer outputs $H = \{h_1, h_2, \ldots, h_n\}$

$$H = \text{BERT}(w_1, w_2, \ldots, w_n), \tag{1}$$

where $h_i \in \mathbb{R}^{1\times d}$ and d is the dimension of BERT layer features.

 https://doi.org/10.1007/s44196-025-00900-w

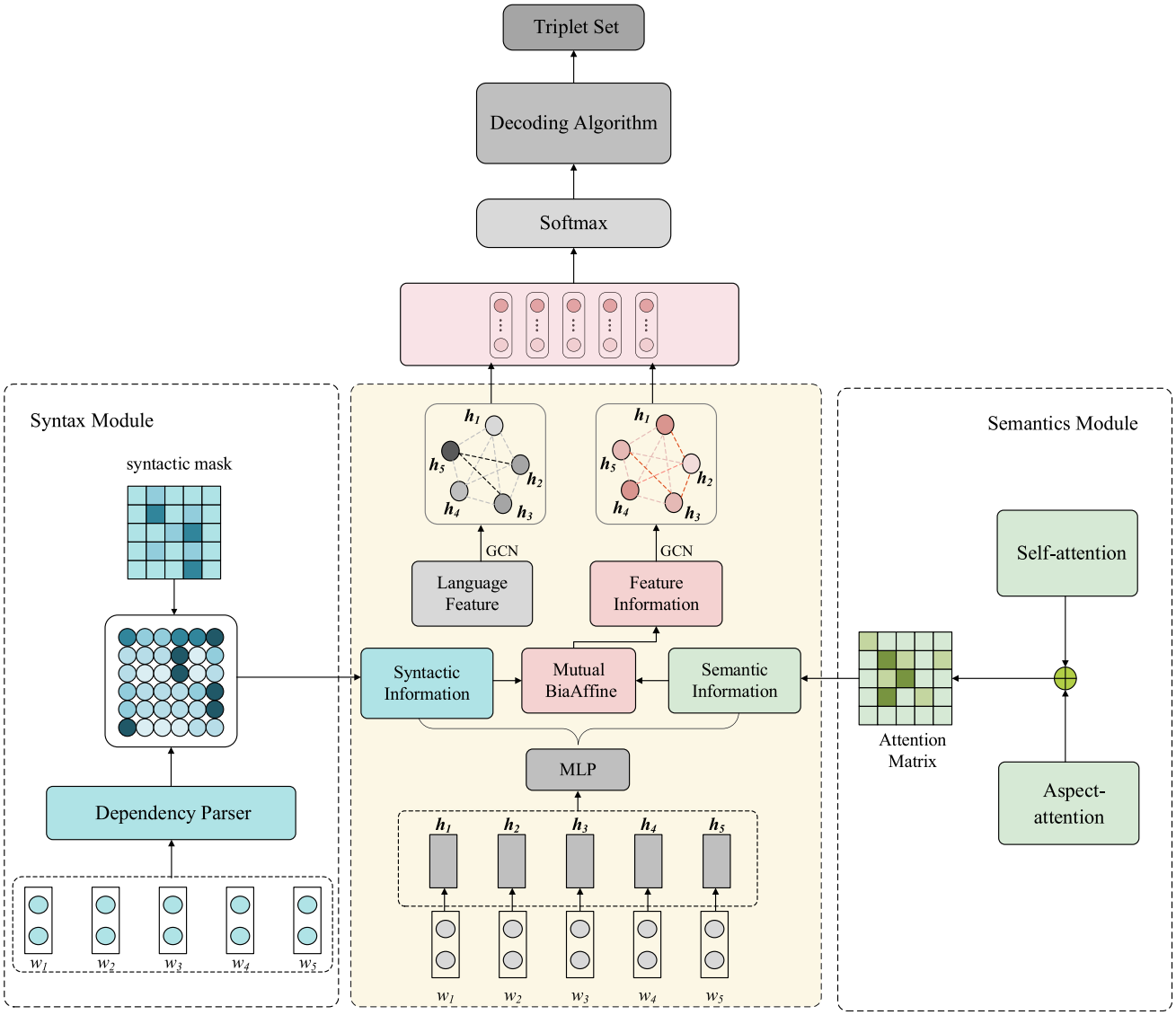


Fig. 2 SSGCN model

3.2 Semantic Information

Self-attention The attention mechanism allows neural networks to better capture the relationships between words in a sentence [38]. It works by calculating attention weights between elements in a sequence, enabling the model to focus on the most relevant parts of the input. In our model, we use the attention score matrix as a semantic representation, as it encodes the relationships between all word pairs in the sentence. By applying the self-attention mechanism, we can effectively capture contextual dependencies across the entire sentence. The detailed calculation process is provided in Eq. (2)

$$A_{\text{self}} = \text{softmax}\left(\frac{QW^{Q} \times ((KW^{K})^{T})}{\sqrt{d_1}}\right), \tag{2}$$

where the matrices Q and K are node features, and W^{Q} and W^{K} are learnable weight matrices. In addition, d_1 is the dimension of the input node features.

Aspect-attention Unlike the self-attention mechanism, which focuses on capturing contextual dependencies for all words in a sentence, the aspect-attention mechanism [39] specifically targets the contextual dependencies of aspect terms. This enables the model to better capture the sentiment polarity of these aspects, thus improving its performance on aspect-specific tasks. The calculation process is provided in Eq. (3) [39]

$$A_{\text{aspect}} = \tanh(H_a W^a \times (K W^k)^T + b), \tag{3}$$

where K is obtained through the input encoding layer, and H_a is aspect term in the sentence. $W^a \in \mathbb{R}^{d\times d}$ and $W^k \in \mathbb{R}^{d\times d}$ are learnable weights.

By combining the self-attention score matrix with the aspect-attention score matrix, the model effectively captures both the global semantic context of the sentence and the aspect-specific semantic dependencies. This integration allows the SSGCN model to comprehensively leverage semantic information from both general and aspect-focused perspectives

$$\tilde{A}_{\text{sem}} = A_{\text{self}} + A_{\text{aspect}}. \tag{4}$$

3.3 Syntactic Information

The adjacency matrix of the syntactic dependency tree effectively represents the syntactic relationship between words, because the adjacency matrix reflects the bidirectionality and uniqueness of the dependency relationship and is also easy to calculate and combine with other algorithms. To encode syntactic information, we construct the adjacency matrix using all dependency arcs in the syntactic tree, as shown in Eq. (5). The adjacency matrix is then normalized to produce a probability matrix A_{syn}, which better reflects the relative syntactic importance of connections

$$\tilde{A}_{\text{syn}} = \begin{cases} 0, & \text{if there is no edge between } w_i \text{ and } w_j \\ 1, & \text{if there is an edge between } w_i \text{ and } w_j, \end{cases} \tag{5}$$

where w_i and w_j represent the ith and jth words in a sentence. If there is an edge between w_i and w_j, the corresponding entry in the adjacency matrix is 1; otherwise, it is 0.

Finally, the syntactic mask matrix is introduced, and the fully connected graph is masked according to different syntactic distances. We define the shortest path between nodes as the syntactic distance and set the syntactic mask matrix with the same number of attention heads. The model can learn local information through smaller syntactic distances, while larger syntactic distances capture global structural information. The definition of the syntactic mask matrix can be simplified to the following equation:

$$M_k^{ij} = \begin{cases} 0, & \text{if } D(i, j) \le k \\ 1, & \text{otherwise,} \end{cases} \tag{6}$$

where $k \in [1, p]$ is the number of heads of multi-head attention, and $D(i, j)$ represents the shortest path distance between nodes i and j. Mask matrix $M = \{M^1, M^2, \ldots, M^p\}$. The final form of semantic matrix A_{sem} is shown in Eq. (7)

$$A_{\text{sem}} = \text{soft max}\left(\tilde{A}_{\text{sem}} + \boldsymbol{M}\right). \tag{7}$$

3.4 Interactive Biaffine Module

An interactive module is constructed to effectively integrate syntactic information, semantic information, and node feature information. This module facilitates the interaction between these different types of information, ensuring

that they complement each other. The specific computation process for this integration is presented in Eqs. (8)–(13)

$$h_i^a = \text{MLP}_a(h_i) \tag{8}$$
$$h_j^o = \text{MLP}_o(h_j) \tag{9}$$
$$g_{\text{syn}} = A_{\text{syn}} \times h_i^a \tag{10}$$
$$g_{\text{sem}} = A_{\text{sem}} \times h_j^o \tag{11}$$
$$\text{affine} = \text{linear}(g_{syn}) \tag{12}$$
$$R^{mb} = \text{affine} \times g_{\text{sem}}^T \tag{13}$$
$$R^{mb} = \text{linear}((\text{MLP}_a(h_i) \times A_{\text{syn}})) \times (\text{MLP}_o(h_j) \times A_{\text{sem}})^T, \tag{14}$$

where a multi-layer perceptron (MLP) is used for this integration process. h_i^a and $h_j^o \in \mathbb{R}^{n \times d_1}$ represent the feature vectors of aspect terms and opinion terms, respectively. g_{syn} and $g_{\text{sem}} \in \mathbb{R}^{n \times d_1}$ are features obtained by integrating node features with syntactic and semantic information, respectively. affine $\in \mathbb{R}^{n \times d_1}$ and $R \in \mathbb{R}^{n \times n \times m}$ represent the relationship tensor, where m denotes the number of different relationship categories. R^{mb} is the adjacency tensor input to the graph convolution layer.

In this study, an interaction module is constructed to achieve in-depth fusion of syntactic information, semantic information, and node features. Specifically, multi-layer perceptrons MLP_a and MLP_o are used to perform non-linear transformations on the input aspect term feature h_i and opinion term feature h_j, respectively, resulting in h_i^a and h_j^o. Subsequently, the syntax-related matrix A_{syn} is multiplied with h_i^a, and the semantics-related matrix A_{sem} is multiplied with h_j^o to obtain the aspect term feature g_{syn} (fused with syntactic information) and the opinion term feature g_{sem} (fused with semantic information), respectively. A linear transformation is then applied to g_{syn} to obtain affine, which is multiplied by the transpose of g_{sem} to produce the matrix R^{mb}, representing the relationship between aspect terms and opinion terms. This matrix characterizes the degree of association between the two after fusing syntactic and semantic features. Equation (14) generalizes the process described by Eqs. (8)–(13), and ultimately generates the adjacency tensor used as input to the graph convolutional layer.

3.5 Graph Convolutional Network

Inspired by traditional convolutional neural networks and graph embedding, the Graph Convolutional Network (GCN) is a type of convolutional neural network that can operate directly on graphs and exploit their structural information [40]. In the graph convolution module, we utilize multiple adjacency matrices to represent different types of relationships. In addition to the adjacency matrix that incorporates syntactic and semantic information, we also use four other matrices, as shown in Fig. 3: the relative distance matrix (R^{rpd}), the tree-based distance matrix (R^{tbd}), the part-of-speech pairs matrix (R^{psc}), and the syntactic dependency type matrix (R^{sdt}). Although these four matrices supplement the syntactic and semantic information to a certain extent, they only reflect the syntactic and semantic information from a macro-perspective and lack the fusion of the syntactic and semantic information on the node feature vector. Therefore, we supplement the syntactic and semantic information acting on the feature vector in the above module. We connect the five matrices mentioned above

$$R = cat(R^{mb}, R^{rpd}, R^{tbd}, R^{psc}, R^{sdt}), \tag{15}$$

where $R \in \mathbb{R}^{N \times N \times 5}$ is the input state in the graph convolutional network module.

The adjacency tensor R is utilized as the input to the Graph Convolutional Network (GCN), serving as the structural representation of the graph. Within the GCN, node feature information is aggregated by combining the adjacency matrix R with the feature vectors of nodes. This operation enables the model to capture and propagate both local and global dependencies in the graph structure. The specific aggregation and update process is defined

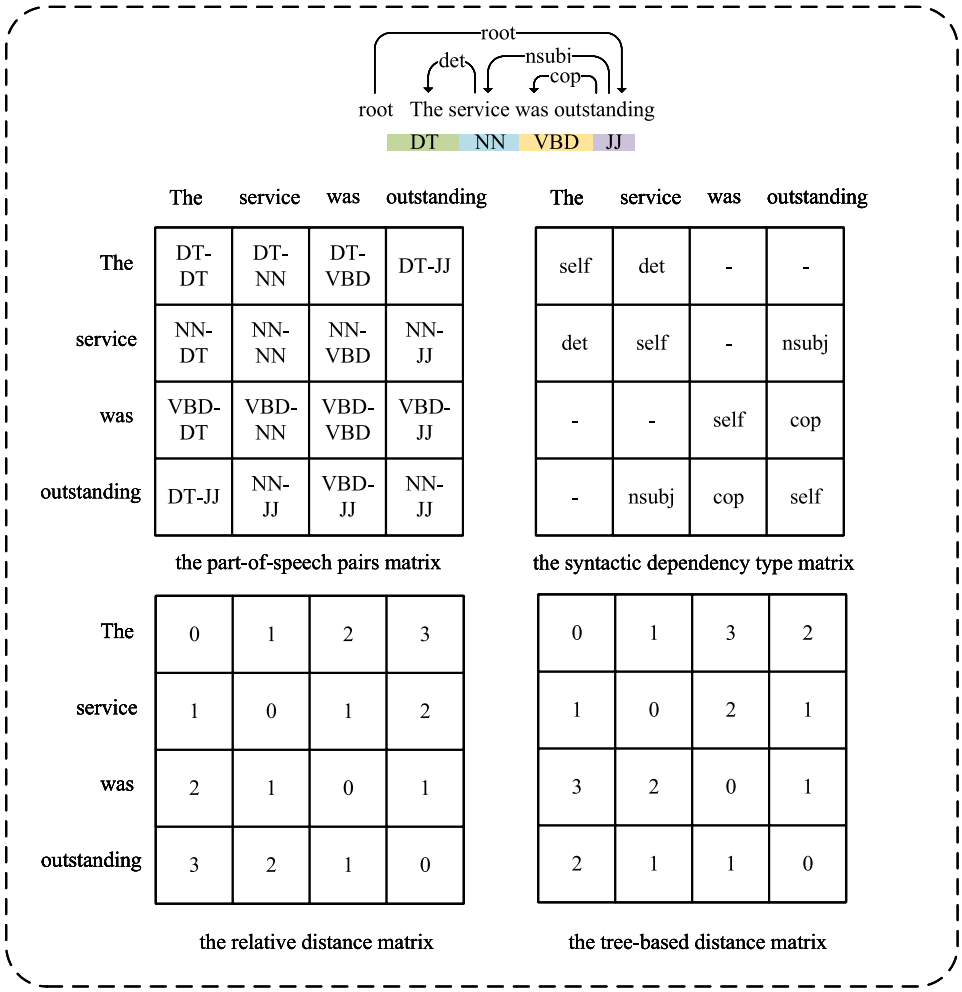


Fig. 3 Four language feature matrices

in Eqs. (16) and (17)

$$O_k = \sigma\left(\sum_{j=1}^{n} R_{ij} W_k^{l+1} H + b_k^{l+1}\right) \tag{16}$$

$$O = f(O_1, O_2, \ldots, O_m), \tag{17}$$

where O represents the output state generated by the graph convolutional network, and $f(.)$ is the average pooling function.

3.6 Sentiment Prediction

In addition, the refining strategy is integrated into the model to combine node and edge representations. For instance, given a word pair (w_i, w_j), the nodes are represented by h_i and h_j, while the edge between the two words is represented by r_{ij}. This strategy enables the model to effectively capture both the individual properties of nodes and the relational information encoded in the edges

$$s_{ij} = h_i \oplus h_j \oplus r_{ij} \oplus r_{ii} \oplus r_{jj}. \tag{18}$$

The output of the linear layer is passed through the softmax function to convert it into a probability distribution p_{ij}, where p_{ij} represents the probability of a specific relationship between the words w_i and w_j. This transformation ensures that the predicted scores are normalized, making them suitable for probabilistic interpretation in subsequent tasks.

$$p_{ij} = softmax(W_s s_{ij} + b_s), \tag{19}$$

where W_s and b_s are the learnable weights and bias terms.

Finally, the prediction results are decoded into triplets. The decoding process begins by iterating over all possible aspect terms and opinion terms in the sentence. For each pair of aspect term and opinion term, the model computes the number of sentiment labels assigned to the tokens between them. Based on this count, the model determines the sentiment polarity. If a sentiment label is successfully identified, a triplet consisting of the aspect term, opinion term, and sentiment label is added to the output list. The detailed decoding procedure is outlined in Algorithm 1.

Algorithm 1 Triplet Decoding for ASTE

Input: tags: sequence of tags representing label attributes, aspect_spans: list of (start, end) indices for aspect terms, opinion_spans: list of (start, end) indices for opinion terms.
Output: $triplets\ T$

```
T = {}
for (al, ar) in aspect_spans do
    for (pl, pl) in opinion_spans do
        tag_num ⇐ {0}
        for i in (al,ar) do
            for j in (ol,or) do
                if al < pl then
                    sentiment lable = tag_num[int(tag[i][j])]
                else
                    sentiment lable = tag_num[int(tag[j][i])]
                end if
            end for
        end for
        if sentiment labels ∈ {POS, NEU, NEG} then
            T ⇐ {al, ar, pl, pr, sentiment}
        end if
    end for
end for
```

3.7 Loss Function

Our training goal is to minimize the total objective function that is as follows:

$$\mathcal{L} = \mathcal{L}_p + \alpha \mathcal{L}_{ba} + \beta(\mathcal{L}_{psc} + \mathcal{L}_{dep} + \mathcal{L}_{tbd} + \mathcal{L}_{rpd}), \tag{20}$$

where the coefficients α and β are utilized to modify the impact of correspondence constraint loss. In the ASTE task, we apply $\mathcal{L}_p$, the typical cross-entropy loss function

$$\mathcal{L}_p = -\sum_{i}^{n}\sum_{j}^{n}\sum_{c\in\mathcal{C}}\prod(y_{ij} = c)log(p_{i,j|c}). \tag{21}$$

Table 1 Dataset statistics

Dataset	res14		lap14		res15		res16	
	#S	#T	#S	#T	#S	#T	#S	#T
Train	1259	2356	899	1452	603	1038	863	1421
Dev	315	580	225	383	151	239	216	348
Test	493	1008	332	547	325	493	328	525

4 Experiment

4.1 Dataset

We evaluate the performance of the SSGCN model using a publicly available dataset. This dataset is an improved version proposed by Wu et al. [28], which builds upon version by Fan et al. [41] and the SemEval Challenge dataset [42–44]. The original SemEval dataset includes annotations for aspect terms and sentiment polarities. Wu et al. further enhanced the dataset by adding annotations for opinion terms, making it more suitable for tasks involving fine-grained opinion extraction. Although more recent datasets have been released by SemEval, we adopt the SemEval-2020 dataset in our experiments to ensure a fair and direct comparison with prior state-of-the-art methods, most of which are benchmarked on this dataset. Moreover, the SemEval-2020 dataset remains one of the most widely used and challenging benchmarks for sentiment triplet extraction and related tasks, providing a solid foundation for evaluating model effectiveness and generalizability. Detailed statistics of the dataset are presented in Table 1, where "#S" and "#T" denote the number of sentences and opinion triplets, respectively.

4.2 Baselines

We compare the SSGCN model with state-of-the-art baseline models:

- **CMLA+** [1], **RINANTE+** [1], and **Li-unified-R+** [1] are pipeline methods, which are extended versions of CMLA [45], RINANTE [46], and Li-unified-R [20] by Peng et al.
- **GTS-BERT** [28] designed a grid tagging scheme and used BERT to learn text features.
- **BMRC** [35] transformed the sentiment triplet extraction task into a multi-round machine reading comprehension task and proposed a bidirectional machine reading comprehension framework.
- **OTE-MTL** [34] extracted relevant aspect terms and sentiment pairs from sentences through a unified generation framework, mainly based on semantic features.
- **JET** [27] proposed the first end-to-end model with a novel position-aware tagging scheme that can jointly extract triples.
- **SA-Transformers** [29] proposed a syntax-aware Transformer model that extended the GCN strategy by making full use of the dependency type of the edge to prevent inappropriate propagation.
- **EMCGCN** [36] adopted multi-channel GCN to process different types of feature information, such as syntactic features in dependency trees and semantic features in context.

By comparing the above models, we aim to evaluate the performance of the SSGCN model in sentiment analysis tasks and explore its advantages and applicability in multi-task learning environments.

4.3 Implementation Details

In the experiments, we employ BERT-base-uncased as the sentence encoder due to its widespread adoption in natural language processing tasks and its pre-trained knowledge encompassing both syntactic and semantic information. The AdamW optimizer [47] is chosen for its effectiveness in handling sparse gradients and preventing

overfitting through weight decay regularization, aligning with recommendations for fine-tuning large pre-trained models. We set the fine-tuning learning rate for BERT to 2×10^{-5}. Lower learning rates led to extremely slow convergence, while higher rates caused the model to diverge. For other trainable parameters, a higher learning rate of 10^{-3} is selected to facilitate faster convergence, as these modules lack the extensive pre-training of BERT. To enhance the model's generalization ability, we apply a dropout rate of 0.4. Higher dropout rates (e.g., 0.5) led to underfitting, while lower dropout rates (e.g., 0.2) resulted in overfitting.

At the architectural level, we configure the hidden state dimensions of BERT and GCN to 768 and 300, respectively, to maintain compatibility with BERT's native embedding size and to balance expressiveness with computational efficiency. These dimensions align with common practices in graph neural networks for NLP [36]. The SSGCN model is trained for 100 epochs with a batch size of 16, a value chosen to balance computational efficiency and model performance. Larger batch sizes (e.g., 32) caused memory issues, while smaller batch sizes (e.g., 8) resulted in noisy gradient updates. For relational constraints, we adjust the hyperparameters α and β, setting them to 0.1 and 0.01, respectively, to regulate their impact during training. Excessively high values of α over-prioritized syntactic constraints, while excessively low values weakened their influence. Similarly, β is set to a smaller value to moderately penalize semantic inconsistencies.

During data preprocessing, we employ the Stanza parser [48] to extract unified syntactic and semantic representations of input sentences, due to its state-of-the-art performance on English tasks and its ability to generate dependency graphs compatible with our GCN architecture. These parameter settings and preprocessing steps are carefully chosen to optimize model performance and ensure the reproducibility of experimental results.

4.4 Main Results

The experimental results are shown in Table 2. Among them, "P" stands for Precision, which measures the proportion of samples predicted as positive by the model that are actually positive, reflecting the accuracy of the model's predictions. "R" stands for Recall, which indicates the proportion of actual positive samples that are correctly identified by the model, reflecting the model's ability to capture positive instances. The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure of the model's precision and recall. The F1-score is the harmonic mean of Precision and Recall, providing a balanced measure of the model's precision and recall. Under the F1-score metric, our proposed SSGCN model outperforms other benchmark experiments on each dataset. Pipeline methods, while effective in some cases, often exhibit inconsistent performance across datasets with varying complexity. This is primarily due to error propagation inherent in their sequential nature. In contrast, models that utilize semantic information perform significantly better overall. Joint learning approaches, such as OTE-MTL and BMRC, leverage multi-task learning and bidirectional mechanisms, respectively, to better capture contextual and semantic dependencies. For example, BMRC achieves an F1-score of 67.49% on the res16 dataset, highlighting its ability to process tasks requiring in-depth contextual understanding. However, its performance on other datasets is relatively lower, possibly due to challenges in capturing long-distance dependencies effectively.

Among methods leveraging syntactic information, models like JET and SA-Transformer demonstrate their advantages in handling long-range dependencies. JET achieves an F1-score of 63.21% on the res16 dataset, reflecting its capability to capture complex syntactic structures. SA-Transformer further improves on this by incorporating syntactic structure with a self-attention mechanism, achieving better results on multiple datasets. GTS-BERT, which combines the pre-trained language model BERT with syntactic information, attains an F1-score of 70.20% on the res14 dataset. However, its performance on the lap14 dataset drops to 54.58%, possibly due to high proportion of short sentences, where syntactic features may not be fully utilized.

In comparison, SSGCN exhibits balanced and superior performance across all datasets. By effectively integrating syntactic and semantic information into node and edge representations, SSGCN achieves deeper feature extraction and captures both local and global dependencies. This capability allows it to handle diverse tasks with varying syntactic and semantic complexity. Unlike models such as EMCGCN, which primarily focus on graph structures,

Table 2 Experimental results (%)

Model		res14			res15			res16			lap14		
		P	R	F1	P	R	F1	P	R	F1	P	R	F1
Pipeline methods	RINANTE+	31.07	37.63	34.03	29.40	26.90	28.00	27.10	20.50	23.30	49.2	46.23	47.68
	CMLA+	40.11	46.63	43.12	34.40	37.60	35.90	43.60	39.80	41.60	31.40	34.60	32.90
	Li-unified-R+	41.44	68.79	51.68	43.34	50.73	46.69	38.19	53.47	44.51	42.25	42.78	42.47
Methods with semantic information	OTE-MTL	64.54	55.57	59.67	58.16	54.02	55.83	48.17	42.43	45.05	54.18	45.20	48.97
	BMRC	–	–	70.01	–	–	58.74	–	–	67.49	–	–	57.83
	EMCGCN	71.85	72.12	71.98	59.89	61.05	60.38	65.08	71.66	68.18	61.46	55.56	58.32
	SSGCN	74.63	71.42	**73.05**	58.28	65.78	**61.74**	67.59	69.55	**68.57**	68.56	52.43	**59.74**
Methods with syntactic information	JET	61.50	55.13	58.14	64.37	44.33	52.50	70.94	57.00	64.21	53.03	33.89	41.35
	SA-Trans former	70.76	65.85	68.22	62.82	58.31	60.48	72.01	62.87	67.13	61.28	48.98	54.44
	GTS-BERT	70.92	69.49	70.20	59.29	58.07	58.67	68.58	66.60	67.58	57.52	51.92	54.58
	EMCGCN	71.85	72.12	71.98	59.98	61.05	60.38	65.08	71.66	68.18	61.46	55.56	58.32
	SSGCN	74.63	71.42	**73.05**	58.28	65.78	**61.74**	67.59	69.55	**68.57**	68.56	52.43	**59.74**

The bold font SSGCN represents the model we proposed

 https://doi.org/10.1007/s44196-025-00900-w

Table 3 Experimental results of ablation study (%)

Model	res14	lap14	res15	res16
W/o SA	75.15	59.27	59.31	68.43
W/o AA	73.49	59.92	63.32	67.56
W/o SAM	73.52	59.88	58.81	71.38
W/o SAAA	70.93	60.04	60.04	66.31

SSGCN achieves a more holistic understanding of sentence structures, leading to consistently high performance across domains.

4.5 Ablation Study

Ablation studies were conducted to evaluate the contribution of self-attention (SA), aspect-attention (AA), the syntactic adjacency matrix (SAM), and their combination (SAAA). Results are shown in Table 3. Removing the self-attention mechanism caused F1-score drops of 0.8%, 4%, and 0.2% on lap14, res15, and res16, respectively. The significant drop on res15 highlights its importance in capturing global semantic relationships and long-range dependencies, especially in complex syntactic structures. Excluding the aspect-attention mechanism resulted in a 1.5% F1 drop on res16, showing its critical role in capturing aspect-specific semantic nuances and handling multi-faceted sentiments. The syntactic adjacency matrix (SAM) also proved essential, with a 4.7% drop on res15 when removed, underscoring the value of syntactic structure in identifying dependencies and phrase relationships. Finally, removing both SA and AA (SAAA) led to F1-score reductions of 3%, 2.8%, and 3.3% on res14, res15, and res16, demonstrating the importance of integrating global semantic and aspect-specific information. In summary, all components significantly contribute to model performance. Self-attention captures global semantics, aspect-attention extracts fine-grained context, and SAM leverages syntactic dependencies, ensuring balanced and robust results across datasets.

4.6 Feature Visualization

To further analyze the role of our proposed module in the ASTE task, we visualize the adjacency tensor that integrates syntactic and semantic features. In Fig. 4, the column coordinates represent the BERT embedding representations of the words, while the row coordinates represent the corresponding labels. These labels include B-A (beginnings of aspect terms), I-A (interiors of aspect terms), A (aspect terms), B-O (beginnings of opinion terms), I-O (interiors of opinion terms), O (opinion terms), POS (positive sentiment polarity of aspect terms), NEU (neutral sentiment polarity of aspect terms), NEG (negative sentiment polarity of aspect terms), and $\perp$ (no label). In fact, the color intensity reflects the magnitude of the probability value. The darker the color, the higher the corresponding probability. In Fig. 4, the darkest color in the "Delicate" row corresponds to O, indicating that "Delicate" is likely to be an opinion term. However, there may be inaccuracies in the labeling of other words, because the label of a word is influenced not only by this tensor matrix but also by other feature matrices.

4.7 Case Study

In Fig. 5, we provide a visual case study illustrating the identification of aspect and opinion terms, with blue indicating aspect terms and orange indicating opinion terms. The green line denotes the alignment of aspect and opinion terms, resulting in a ternary with positive affective polarity. The opinion term "accommodating" is difficult for GTS-BERT and BMRC to identify, while all methods correctly predicted "courteous". This discrepancy occurs, because "accommodating" is further away from "wait staff" than "courteous". Therefore, these methods overlook the triplet ("wait staff", "accommodating", positive), while our SSGCN could accurately capture it. We believe

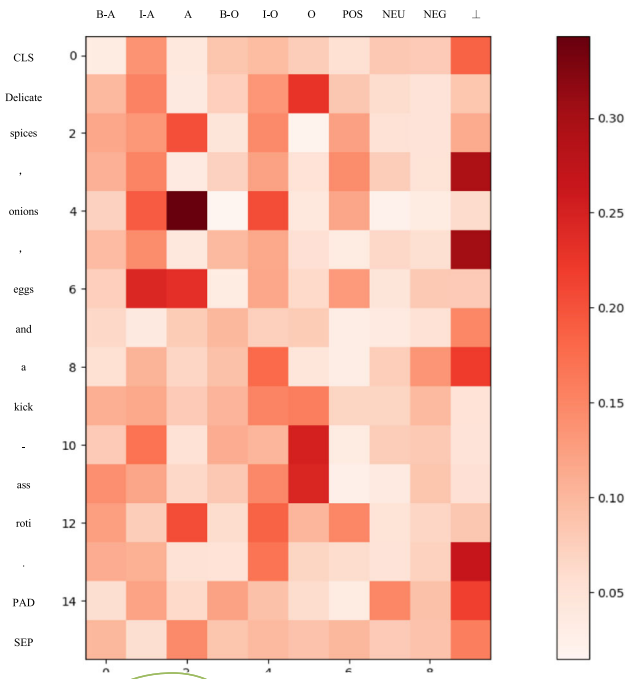


Fig. 4 Visualization of adjacency tensor

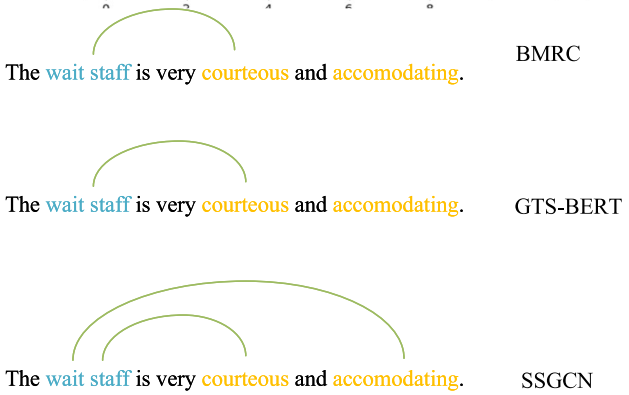


Fig. 5 Case studies of our SSGCN model compared with different baselines

that the crucial factor lies in the substantial connection between "accommodating" and "wait staff" facilitated by semantic and syntactic features.

5 Conclusions

This paper proposes a Syntax-Semantics Graph Convolutional Network (SSGCN) model for the Aspect Sentiment Triplet Extraction (ASTE) task, aiming to effectively integrate syntactic and semantic information within sentences. The SSGCN model consists of three main components: a syntax module, a semantics module, and an interaction module. In the semantics module, we incorporate self-attention and aspect-attention mechanisms to comprehensively capture both global contextual dependencies and aspect-specific contextual dependencies. This design enables the model to not only extract the overall semantic structure of a sentence but also focus on fine-grained semantic information related to specific aspect terms. By leveraging these mechanisms, the model achieves a more precise understanding of semantic relationships. The syntax module utilizes adjacency matrices generated from dependency parsers to extract rich syntactic information. Additionally, we construct mask matrices based on word distances to refine the syntactic representation and emphasize key relationships between words in the sentence. This ensures that the syntactic structure is fully utilized to support triplet extraction. Finally, the interaction module integrates the outputs of the syntax and semantics modules, enabling complementary interactions between syntactic and semantic features. This fusion ensures a balanced and robust representation for the ASTE task. We

conducted extensive experiments on public datasets, and the results demonstrate that the proposed SSGCN model achieves significant improvements in F1 scores compared to baseline models. This performance gain highlights the effectiveness of integrating syntactic and semantic information, particularly in addressing challenges such as long-distance dependencies and polysemy ambiguity in aspect-based sentiment analysis.

Author Contributions J.Z.: Primarily responsible for the research design, experimental setup, and data analysis. This includes formulating the research problem, selecting and implementing methods, organizing and conducting experiments, as well as analyzing and interpreting the results. S.X.: Oversaw the overall conceptualization of the paper, guided the experimental design, determined the data organization and analysis framework, conducted the final review and submission of the manuscript, and ensured proper distribution of responsibilities and recognition among the authors. X.G.: Assisted in designing the research methodology, collected and preprocessed experimental data, participated in the implementation and fine-tuning of the algorithm models, assisted in analyzing the experimental results, and contributed to the writing and revision of the manuscript. Z.T.: Responsible for writing the literature review, providing theoretical support and background analysis, assisting in the discussion of methods, contributing to the experimental validation section, and offering suggestions for structuring and refining the paper.

Funding No funding.

Data Availability The data that support the findings of this study are openly available at https://github.com/NJUNLP/GTS from Wu et al.'s paper [28].

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Peng, H., Xu, L., Bing, L., Huang, F., Lu, W., Si, L.: Knowing what, how and why: a near complete solution for aspect-based sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34. USA, New York, pp. 8600–8607 (2020)
2. Li, R., Chen, H., Feng, F., Ma, Z., Wang, X., Hovy, E.: Dual graph convolutional networks 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 (Volume 1: Long Papers), pp. 6319–6329 (2021)
3. Machado, M.T., Pardo, T.A.S.: Evaluating methods for extraction of aspect terms in opinion texts in Portuguese: the challenges of implicit aspects. In: The 13th Conference on Language Resources and Evaluation, France, Marseille, pp. 3819–3828 (2022)
4. Wang, Q., Wen, Z., Zhao, Q., Yang, M., Xu, R.: Progressive self-training with discriminator for aspect term extraction. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 257–268 (2021)
5. Xu, H., Liu, B., Shu, L., Philip, S.Y.: Double embeddings and cnn-based sequence labeling for aspect extraction. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Australia, Melbourne, pp. 592–598 (2018)
6. Li, X., Lam, W.: Deep multi-task learning for aspect term extraction with memory interaction. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Denmark, Copenhagen, pp. 2886–2892 (2017)
7. Feng, Y., Rao, Y., Tang, Y., Wang, N., Liu, H.: Target-specified sequence labeling with multi-head self-attention for target-oriented opinion words extraction. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1805–1815 (2021)

8. Li, Y., Wang, F., Zhong, S.-H.: Training entire-space models for target-oriented opinion words extraction. In: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Spain, Madrid, pp. 1875–1879 (2022)
9. Mensah, S., Sun, K., Aletras, N.: An empirical study on leveraging position embeddings for target-oriented opinion words extraction. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 9174–9179 (2021)
10. Wang, B., Shen, T., Long, G., Zhou, T., Chang, Y.: Eliminating sentiment bias for aspect-level sentiment classification with unsupervised opinion extraction. In: Findings of the Association for Computational Linguistics: EMNLP 2021, Dominican Republic, Punta Cana, pp. 3002–3012 (2021)
11. Hou, X., Qi, P., Wang, G., Ying, R., Huang, J., He, X., Zhou, B.: Graph ensemble learning over multiple dependency trees for aspect-level sentiment classification. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 2884–2894 (2021)
12. Tian, Y., Chen, G., Song, Y.: Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 2910–2922 (2021)
13. Zhou, Y., Liao, L., Gao, Y., Jie, Z., Lu, W.: To be closer: learning to link up aspects with opinions. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 3899–3909 (2021)
14. Wang, S., Mazumder, S., Liu, B., Zhou, M., Chang, Y.: Target-sensitive memory networks for aspect sentiment classification. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Australia, Melbourne, pp. 957–967 (2018)
15. Wu, S., Fei, H., Ren, Y., Ji, D., Li, J.: Learn from syntax: improving pair-wise aspect and opinion terms extraction with rich syntactic knowledge. In: Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21) (2021)
16. Chakraborty, A.: Aspect and opinion term extraction using graph attention network. In: Proceedings of the 20th International Conference on Natural Language Processing (ICON), India, Goa, pp. 594–602 (2023)
17. Gao, L., Wang, Y., Liu, T., Wang, J., Zhang, L., Liao, J.: Question-driven span labeling model for aspect—opinion pair extraction. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 12875–12883 (2021)
18. Yu, G., Li, J., Luo, L., Meng, Y., Ao, X., He, Q.: Self question-answering: aspect-based sentiment analysis by role flipped machine reading comprehension. In: Findings of the Association for Computational Linguistics: EMNLP 2021, Dominican Republic, Punta Cana, pp. 1331–1342 (2021)
19. Li, H., Lu, W.: Learning latent sentiment scopes for entity-level sentiment analysis. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31 (2017)
20. Li, X., Bing, L., Li, P., Lam, W.: A unified model for opinion target extraction and target sentiment prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33. USA, pp. 6714–6721 (2019)
21. Liu, J., Teng, Z., Cui, L., Liu, H., Zhang, Y.: Solving aspect category sentiment analysis as a text generation task. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, pp. 4406–4416 (2021)
22. Cai, H., Tu, Y., Zhou, X., Yu, J., Xia, R.: Aspect-category based sentiment analysis with hierarchical graph convolutional network. In: Proceedings of the 28th International Conference on Computational Linguistics, pp. 833–843 (2020)
23. Dai, J., Yan, H., Sun, T., Liu, P., Qiu, X.: Does syntax matter? A strong baseline for aspect-based sentiment analysis with RoBERTa. In: Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1816–1829 (2021)
24. Ye, H., Yan, Z., Luo, Z., Chao, W.: Dependency-tree based convolutional neural networks for aspect term extraction. In: Advances in Knowledge Discovery and Data Mining: 21st Pacific-Asia Conference, PAKDD 2017, Jeju, South Korea, May 23–26, 2017, Proceedings, Part II 21, South Korea, Jeju, pp. 350–362 (2017)
25. Feng, L., Zeng, B., He, L., Xu, M., Deng, H., Chen, P., Huang, Z., Du, W.: Improving span-based aspect sentiment triplet extraction with abundant syntax knowledge. Neural Process. Lett. **55**(5), 5833–5854 (2023)
26. Wang, G., Wang, B., Xu, F., Wang, R., Zhu, Z., Liu, P.: Domain-consistent syntactic representation for cross-domain aspect sentiment triplet extraction. Expert Syst. Appl. **256**, 124854 (2024)
27. Xu, L., Li, H., Lu, W., Bing, L.: Position-aware tagging for aspect sentiment triplet extraction. In: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 2339–2349 (2020)
28. Wu, Z., Ying, C., Zhao, F., Fan, Z., Dai, X., Xia, R.: Grid tagging scheme for aspect-oriented fine-grained opinion extraction. In: Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 2576–2585 (2020)
29. Yuan, L., Wang, J., Yu, L.-C., Zhang, X.: Encoding syntactic information into transformers for aspect-based sentiment triplet extraction. IEEE Trans. Affect. Comput. **15**(2), 722–735 (2023)
30. Xu, L., Chia, Y.K., Bing, L.: 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 (Volume 1: Long Papers), pp. 4755–4766 (2021)
31. Chen, Y., Keming, C., Sun, X., Zhang, Z.: A span-level bidirectional network for aspect sentiment triplet extraction. In: Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, United Arab Emirates, Abu Dhabi, pp. 4300–4309 (2022)

32. Chen, Y., Zhang, Z., Zhou, G., Sun, X., Chen, K.: Span-based dual-decoder framework for aspect sentiment triplet extraction. Neurocomputing **492**, 211–221 (2022)
33. Jiang, B., Liang, S., Liu, P., Dong, K., Li, H.: A semantically enhanced dual encoder for aspect sentiment triplet extraction. Neurocomputing **562**, 126917 (2023)
34. Zhang, C., Li, Q., Song, D., Wang, B.: A multi-task learning framework for opinion triplet extraction. In: Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 819–828 (2020)
35. Chen, S., Wang, Y., Liu, J., Wang, Y.: Bidirectional machine reading comprehension for aspect sentiment triplet extraction. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 12666–12674 (2021)
36. Chen, H., Zhai, Z., Feng, F., Li, R., Wang, X.: 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), Ireland, Dublin, pp. 2974–2985 (2022)
37. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the Conference on the North American Chapter of the Association for Computational Linguistics, Minnesota, Minneapolis, pp. 4171–4186 (2019)
38. Vaswani, A., Shazeer, N., Parmar, N.: Attention is all you need. In: Roceedings of the 31st International Conference on Neural Information Processing Systems, USA, California, pp. 6000–6010 (2017)
39. Zhang, Z., Zhou, Z., Wang, Y.: SSEGCN: syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis. In: Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, United States, Seattle, pp. 4916–4925 (2022)
40. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks (2017)
41. Fan, Z., Wu, Z., Dai, X., Huang, S., Chen, J.: 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, Volume 1 (Long and Short Papers), Minnesota, Minneapolis, pp. 2509–2518 (2019)
42. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O.: SemEval-2016 task 5: aspect based sentiment analysis. In: Proceedings of the 10th International Workshop on Semantic Evaluation, USA, San Diego, pp. 19–30 (2016)
43. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O.: SemEval-2015 task 12: aspect based sentiment analysis. In: Proceedings of the 9th International Workshop on Semantic Evaluation, USA, Denver, pp. 486–495 (2015)
44. Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O.: SemEval-2014 task 4: aspect based sentiment analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation, Ireland, Dublin, pp. 27–35 (2014)
45. Wang, W., Pan, S.J., Dahlmeier, D., Xiao, X.: Coupled multi-layer attentions for co-extraction of aspect and opinion terms. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31. USA, California (2017)
46. Dai, H., Song, Y.: Neural aspect and opinion term extraction with mined rules as weak supervision. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Italy, Florence, pp. 5268–5277 (2019)
47. Loshchilov, I., Hutter, F.: Fixing weight decay regularization in ADAM. In: Proceedings of the Sixth International Conference on Learning Representations, Canada, Vancouver (2018)
48. Qi, P., Zhang, Y., Zhang, Y., Bolton, J., Manning, C.D.: Stanza: a python natural language processing toolkit for many human languages. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 101–108 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Jingyun Zhang[1,2] · Shuwei Xu[1,2] · Xin Gao[1,2] · Zhiwei Tang[1,2]

✉ Shuwei Xu
xsw@henu.edu.cn

Jingyun Zhang
zjy056@henu.edu.cn

Xin Gao
gaoxin@henu.edu.cn

Zhiwei Tang
tzwlearning@henu.edu.cn

[1] School of Software, Henan University, Kaifeng 475000, China

[2] Henan Provincial Engineering Research Center of Intelligent Data Processing, Henan University, Kaifeng 475000, China