Neurocomputing 507 (2022) 315–324

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Dependency graph enhanced interactive attention network for aspect sentiment triplet extraction

Lingling Shi [a], Donghong Han [a,b,*], Jiayi Han [c], Baiyou Qiao [a], Gang Wu [a]

[a] School of Computer Science and Engineering, Northeastern University, Shenyang, China
[b] Key Laboratory of Intelligent Computing in Medical Image of Ministry of Education, Northeastern University, Shenyang, China
[c] Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China

ARTICLE INFO

Article history:
Received 23 April 2022
Revised 26 June 2022
Accepted 24 July 2022
Available online 10 August 2022
Communicated by Zidong Wang

Keyword:
Interactive attention mechanism
Part-of-Speech
Dependency graph
Aspect sentiment triplet extraction

ABSTRACT

Aspect sentiment triplet extraction is an extremely daunting task designed to identify the triplets from comments, where each triplet is composed of an aspect term, the related opinion term, and the sentiment between them. Existing research efforts majorly construct a novel tagging scheme to avoid the disadvantages of pipeline methods. However, the improvement is limited due to neglecting the implicit grammatical relationships among the three elements in a triplet. To cope with this limitation, we put forward an innovative Dependency Graph Enhanced Interactive Attention Network, which explicitly introduces the syntactic and semantic relationships between words. Specifically, an interactive attention mechanism is conceived to jointly consider both the contextual features learned from Bi-directional Long Short-Term Memory and the syntactic dependencies learned from the correspondent dependency graph in an iterative interaction manner. In addition, we notice that words with different Part-of-Speech categories have different contributions to the semantic expression of sentences. Accordingly, the information of different Part-of-Speech categories is recognized during the modeling process to properly capture the semantic relationships. Experiments on the benchmark datasets originally derived from SemEval Challenges illustrate that our presented approach has superiority over strong baselines.

1. Introduction

Aspect Sentiment Triplet Extraction (ASTE) [1] is committed to identifying the triplets from a given comment, containing aspect terms, related opinion terms, and the sentiment of every aspect term. As shown in Fig. 1, where the comment is from the laptop domain, the triplets of this comment are (*memory*, *bad*, *negative*) and (*battery life*, *great*, *positive*), respectively.

ASTE is a finer-grained work of Aspect Based Sentiment Analysis (ABSA) [2]. ABSA requires extracting aspect terms in a comment as well as the sentiment associated with each aspect term, which usually contains the following basic subtasks: Aspect Terms Extraction (ATE), Opinion Terms Extraction (OTE), and Aspect-based Sentiment Classification (ASC). ABSA has been receiving extensive attention, and the existing works can be broadly classified as follows: (1) The single basic subtask: ATE [3–5] focuses on extracting aspect terms, while OTE [6–8] is devoted to extracting the opinion terms that corresponded to aspect terms. ASC [9–13] is proposed to judge the sentiment of the given aspect term. (2) The combination of two basic subtasks: AESC [14–16] is the combination of ATE and ASC to handle them simultaneously, while Pair [17] is the combination of ATE and OTE, which aims to jointly recognize aspect terms and related opinion terms. (3) The combination of all basic subtasks: ASTE is first mentioned in [1] and a two-stage pipeline approach is proposed but easily suffers from error propagation. To avoid the disadvantages of the pipeline approaches, the authors of [18] propose a method stemming from a novel position-aware tagging scheme to generate final triplets at the same time. Similarly, a brand-new Grid Tagging Scheme (GTS) [19] is presented with the purpose of jointly extracting the triplets. However, these methods focus on formalizing ASTE as a unified task and fail to effectively capture the abundant interactions among triplet elements.

One important observation is that various relationships exist among the three elements in a triplet, such as syntactic and semantic relationships. These relationships are closely related to triplet extraction, but they have not been seriously taken. As shown in Fig. 1, there is a direct syntactic dependency between aspect term

* Corresponding author at: School of Computer Science and Engineering, Northeastern University, Shenyang, China.

E-mail addresses: 2001801@stu.neu.edu.cn (L. Shi), handonghong@cse.neu.edu.cn (D. Han), 18110850001@fudan.edu.cn (J. Han), qiaobaiyou@cse.neu.edu.cn (B. Qiao), wugang@cse.neu.edu.cn (G. Wu).

https://doi.org/10.1016/j.neucom.2022.07.067
0925-2312/

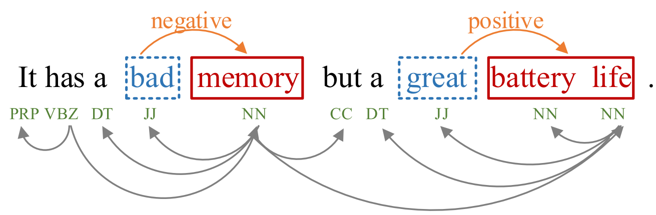

Fig. 1. A typical example of ASTE. The aspect terms are surrounded by solid boxes, while opinion terms are wrapped by dashed boxes. The yellow arrow above indicates these two words can form a valid triplet. The gray arrow below represents these two words have a syntactic dependency, and the green words denote the Part-of-Speech corresponding to each word.

memory and opinion term *bad*, and this dependency also exists between aspect term *battery life* and opinion term *great*. Furthermore, some studies [20,21] conclude that certain Part-of-Speech (POS) categories are highly related to semantic expression. As presented in Fig. 1, aspect terms, *memory* and *battery life*, are nouns, and opinion terms, *bad* and *great*, are adjectives. Intuitively, adverbs are conducive to the prediction of the corresponding sentiment. Meanwhile, verbs, as the main components of sentence structure, can promote the correspondence of aspect terms and opinion terms. Previous achievement [20] can ensure the rationality of these intuitions.

To take full advantages of abundant relationships among the three elements in a triplet, we present a Dependency Graph Enhanced Interactive Attention Network (DGEIAN), which explicitly considers the implicit grammatical relationships between words, including syntactic dependencies, contextual features, and POS information. To be specific, a fresh interactive attention mechanism is formulated to learn syntactic dependencies and contextual features in an iterative interaction manner. The mechanism consists of two self attention layers, followed by two parallel modules, a Bi-directional Long Short-Term Memory (Bi-LSTM) [22] and a multi-layer Graph Convolutional Network (GCN) [23], where contextual and syntactic representations can be reinforced and fused through exchanging their relevant features. The final representations produced by the Bi-LSTM module are then used via GTS for triplet extraction. Moreover, we introduce POS information to promote the learning of semantic information. Not all POS categories can part with effective information for triplet extraction. Therefore, to better leverage the POS information, we distinguish noun, adjective, adverb, and verb categories from other categories during the modeling process.

Different from the existing pipeline methods, our method applies the tagging scheme to extract all triplets in the sentence at one time after obtaining the high-level features of the sentence, which can avoid the influence of propagation errors. Unlike the previous methods that only rely on the tagging scheme, our method benefits from the implicit grammatical relationships in sentences. Specifically, DGEIAN introduces the syntactic dependency tree to learn the syntactic information of sentences and reduces the adverse effects of noise information through the interactive attention mechanism, which can effectively enhance the correspondence between aspect terms and opinion terms and reduce parsing errors. DGEIAN also introduces POS information of words, which can distinguish word categories to accurately identify aspect terms and opinion terms in sentences. To verify the effectiveness of our proposed method, we conduct extensive experiments on several benchmark datasets. The experimental results show that our DGEIAN is superior in the ASTE task.

Our contributions can be summarized as follows:

- We put forward an innovative approach named DGEIAN for the ASTE task to end-to-end settle the imperfections of previous methods. Such an approach can effectively capture the interactions among the triplet elements by explicitly introducing the implicit grammatical relationships between words.
- We conceive a novel interactive attention mechanism to jointly consider the syntactic dependencies and contextual features of sentences, where these two representations can be enhanced via interlacing their relevant features. We also conduct extensive analysis to demonstrate its effectiveness.
- We model the information of certain POS categories to recognize the distinctions between words, which facilitates the DGEIAN in obtaining more abundant semantic relationships. We also investigate the contribution of different combinations of POS categories for triplet extraction through a set of experiments.
- We carry out extensive experiments on the benchmark datasets. The results indicate the superiority of DGEIAN in the ASTE task.

The following sections are arranged as follows. Section 2 exhibits our proposed DGEIAN. Section 3 shows experiment results. Section 4 and Section 5 are respectively related work and conclusions.

2. Our Approach

This section first gives the definition of ASTE, then presents the details of our DGEIAN. The overview of DGEIAN is shown in Fig. 2. For a given input text, we first obtain the word embeddings via employing double embeddings and POS embeddings, then utilize Bi-LSTM and GCN to extract contextual and syntactic representations, respectively. After that, these hidden representations are fed into our proposed interactive attention mechanism, with the corresponding dependency graph. At last, we apply an attention layer to obtain the high-level features and extract triplets via GTS.

2.1. Definition of ASTE

Given a n-word comment sentence $\mathscr{S} = \{w_1, w_2, \ldots, w_n\}$, the purpose of ASTE is identifying a group of triplets $\mathscr{T} = \{(at, ot, s)_m\}_{m=1}^{|T|}$, where $(at, ot, s)_m$ denotes the m-th triplet, and at, ot, and s represent the aspect term, the opinion term, and the sentiment of this triplet respectively. $|T|$ indicates the triplet set length and $s \in \{\text{positive}, \text{negative}, \text{neutral}\}$.

2.2. The DGEIAN Framework

2.2.1. Embedding Layer

High-quality word embedding is instrumental in correctly capturing the semantic information of each word in a sentence. Different words usually have different meanings, which is based on the general domain. The same word usually has different meanings in different contexts, which is based on the specific domain. In order to fully leverage these two meanings, we utilize double embeddings to obtain the initial word embedding $E_w \in \mathscr{R}^{n \times d_w}$ of the sentence with n length, where d_w denotes the dimension of the initial word embedding. Here, double embeddings contain a domain-general embedding and a domain-specific embedding, which are widely applied in the initialization of word embedding.

The observations in Fig. 1 show that certain POS categories are capable of providing effective information in triplet extraction. To better leverage the POS information, we first use NLTK[1] to mark the POS of each word of the sentence. In order to distinguish words

[1] https://www.nltk.org/

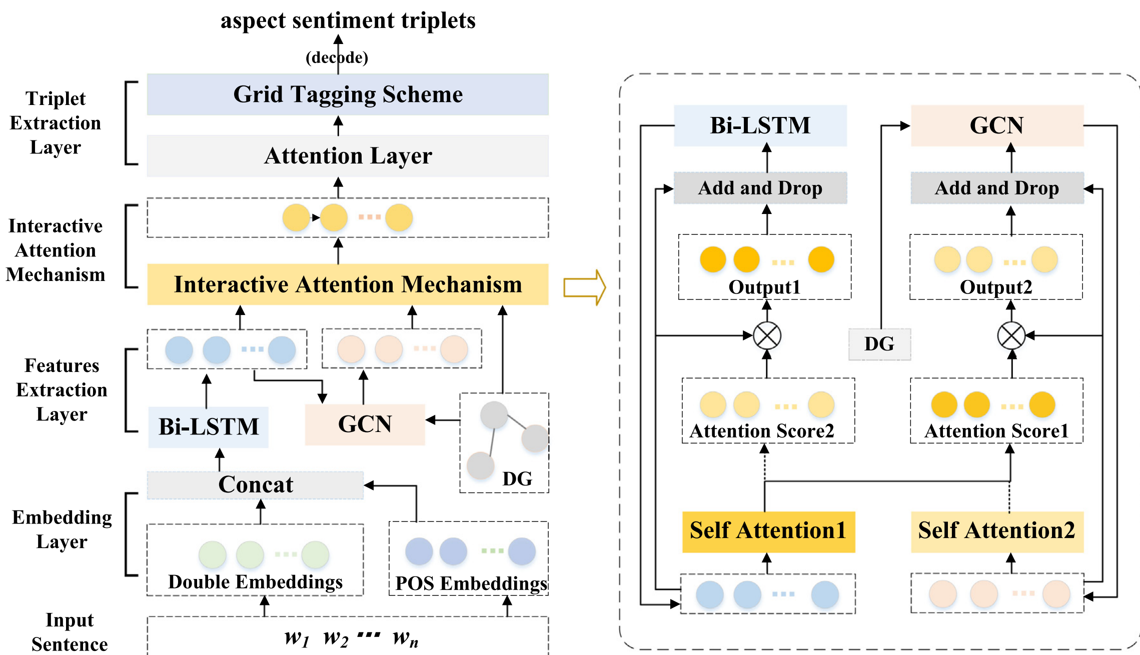

Fig. 2. An overall demonstration of our proposed DGEIAN.

with different POS categories and model their weights in the sentence, we divide POS into five categories to obtain the task-adapted POS sequence $\mathscr{P} = [p_1, p_2, \ldots, p_n], p_i \in \{p_{noun}, p_{adj}, p_{adv}, p_{verb}, p_{others}\}$, where $p_{noun}, p_{adj}, p_{adv}, p_{verb}$ and p_{others} respectively represent noun, adjective, adverb, verb and others categories. For example, given the sentence "It has a bad memory but a great battery life.", the task-adapted POS sequence can be represented as $\mathscr{P} = [p_{others}, p_{verb}, p_{others}, p_{adj}, p_{noun}, p_{others}, p_{others}, p_{adj}, p_{noun}, p_{noun}, p_{others}]$. The corresponding POS embedding $E_p \in \mathscr{R}^{n\times d_p}$ is initialized randomly and updated during the training process, where d_p denotes the dimension of the POS embedding.

To explicitly incorporate POS information into the embedding of each word, we concatenate E_w and E_p as $E \in \mathscr{R}^{n\times(d_w+d_p)}$, which is passed to the next layer to be contextualized.

2.2.2. Features Extraction Layer

As previously motivated, the three elements in a triplet are highly contextual. Bi-LSTM is widely used as the encoder because of its ability to learn contextual features. Hence, we utilize Bi-LSTM to produce the contextual representation $X \in \mathscr{R}^{n\times d_l}$ of the sentence. Here, d_l represents the hidden state dimension of Bi-LSTM. We can simplify the process as follow:

$$X = \mathrm{Bi-LSTM}(E) \tag{1}$$

To capture the syntactic dependencies between words, we employ the dependency graph implemented on a multi-layer GCN, which makes the representation of each word further enriched by the information of its syntactically related words. In our case, the initial features of GCN are the contextual representations, We can express this process as below:

$$H^{(l)} = \mathrm{ReLu}(AH^{(l-1)}W) \tag{2}$$

where $H^{(l)} \in \mathscr{R}^{n\times d_g}$ denotes the output of the l-th GCN layer, $H^{(0)} \in \mathscr{R}^{n\times d_l}$ denotes the initial input of GCN and $H^{(0)} = X$. Here, d_g represents the hidden state dimension of GCN. $A \in \mathscr{R}^{n\times n}$ denotes the adjacency matrix obtained from corresponding dependency graph. We also attach a self-loop for each word. $W \in \mathscr{R}^{d_l\times d_g}$ is a weight matrix that can be updated during training. To simplify the subsequent description, the final output of GCN is defined as H.

Then these two hidden representations and the corresponding dependency graph are fed into our proposed interactive attention mechanism.

2.2.3. Interactive Attention Mechanism

We develop an interactive attention mechanism to enhance the contextual representation and the syntactic representation in an iterative interaction manner. Specifically, as demonstrated in Fig. 2, the interactive attention mechanism comprises two self attention layers, followed by a Bi-LSTM and a multi-layer GCN on top respectively. Note that the Bi-LSTM and multi-layer GCN here are the same as the corresponding networks of Section 2.2.2 in our implementation.

Assume that there are two inputs $S_t^l \in \mathscr{R}^{n\times d_l}$ and $S_t^g \in \mathscr{R}^{n\times d_g}$, where $t \in [1,N]$ represents it is the t-th iteration currently and $S_0^l = X, S_0^g = H$. We apply the interactive attention mechanism to reinforce these two representations.

We first pass these two representations to the self attention mechanisms to obtain the corresponding attention scores α_t^l and α_t^g, which represent the contextual-based and syntactic-based weights of each word in the sentence. In our implementation, the attention scores are formulated as below:

$$\alpha_t^l = \mathrm{softmax}(S_t^l {S_t^l}^\top) \tag{3}$$

$$\alpha_t^g = \mathrm{softmax}(S_t^g {S_t^g}^\top) \tag{4}$$

To fuse contextual and syntactic features, we apply these two attention scores to each other's hidden representations to interlace their relevant features, which means that the attention score α_t^l containing contextual features is fed to syntactic representation S_t^g, while the attention score α_t^g including syntactic features is passed to contextual representation S_t^l. This process can be formulated as:

$$S_t^{l\prime} = \alpha_t^g S_t^l \tag{5}$$

$$S_t^{g'} = \alpha_t^l S_t^g \tag{6}$$

where $S_t^{l'}$ and $S_t^{g'}$ denote the hidden representations fused with each other's features in the t-th iteration.

To effectively alleviate the occurrence of over-fitting, the Dropout function is applied for the above two hidden representations. Subsequently, the higher-level features are extracted as follows:

$$S_{t+1}^l = \mathrm{Bi-LSTM}(\mathrm{Dropout}(S_t^{l'}) + S_t^l) \tag{7}$$

$$S_{t+1}^g = \mathrm{GCN}(\mathrm{Dropout}(S_t^{g'}) + S_t^g) \tag{8}$$

where S_{t+1}^l and S_{t+1}^g represent the output of the interactive attention mechanism in the t-th iteration.

The process denoted by Eqs. (3)–(8) is supported to iterate for N times for adequate feature fusion. We choose the output of Bi-LSTM as the final output of the interactive attention mechanism to avoid introducing too much noise information.

2.2.4. GTS

GTS adopts a unified tagging scheme to extract triplets, which has superior performance in the ASTE task. Hence, we apply this tagging scheme and corresponding decoding algorithm to generate the final triplets in the sentence. Specially, we pass the sentence representation that combines the contextual, syntactic, and POS information to an attention layer, which is the same as the corresponding networks of Section 2.2.3, to extract high-level features before applying GTS.

In GTS, the set $\{A, O, Pos, Neu, Neg, N\}$ is applied to tag the relation of any two words of the sentence. A and O respectively represent these two words are the same aspect term and the same opinion term. When these two words can form a valid triplet, the corresponding sentiment, positive, neutral, or negative, is denoted by the sentiment label Pos, Neu, or Neg. Otherwise, GTS uses N to indicate that there is no above relationship. Fig. 3 exhibits a tagging example. In the decoding algorithm, GTS first determines aspect terms and opinion terms according to the tags on the main diagonal and then judges whether any aspect term and any opinion term can form a valid triplet. If these two terms can form a valid triplet, the most sentiment label is considered the final sentiment of the triplet.

2.2.5. Loss Function

The training goal of DGEIAN is formulated to minimize cross-entropy loss:

$$\mathcal{L} = -\sum_{i=1}^{n}\sum_{j=i}^{n}\sum_{k\in s}\mathbb{I}(y_{ij}=k)\log\left(P_{i,j|k}^L\right) \tag{9}$$

where y_{ij} and $P_{i,j}^L$ denote the ground truth and predicted tagging distribution. s represents the tag set $\{A, O, Pos, Neu, Neg, N\}$.

3. Experiments

3.1. Datasets and Metrics

We mainly carry out experiments on the datasets created in [19]. The detailed information of these datasets is manifested Table 1. These datasets originally come from SemEval Challenges [2,24,25], including one dataset from the laptop domain and three datasets from the restaurant domain. To be fair, we also provide the main results on the ASTE-Data-V2 dataset released in [18], and define the DGEIAN here as DGEIAN_V2.

In order to objectively measure the performances of different methods, we choose the evaluation metrics widely adopt in the ASTE task, including Precision (P), Recall (R), and F1 score (F1). It is considered that a triplet is correct when all elements in this triplet are consistent with their ground truth span.

3.2. Baselines

We select the approaches that are competitive in the ASTE task to demonstrate the performance of DGEIAN. These methods can be classified into pipeline methods and end-to-end methods.

Pipeline methods

- **CMLA+, RINANTE+, and Li-unified-R.** The authors of [26,27] propose two different methods named CMLA and RINANTE to jointly extract aspect terms and opinion terms from comments. Subsequently, a novel method [28] based on the unified tagging scheme is suggested to extract word pairs (aspect term, sentiment). The authors of [1] ameliorate these methods to detect triplets, which are renamed CMLA+, RINANTE+, and Li-unified-R.
- **Peng-unified-R + PD.** A pipeline approach is presented in [1]. The proposed model first respectively predicts all word pairs (aspect term, sentiment) and opinion terms to generate all candidate triplets, then a classifier profiting from the MLP is applied to judge the rationality of each triplet.
- **Peng-unified-R + LOG.** The authors of [19] study a pipeline approach, which first employs the method proposed in [1] and then adapts the method named IOG [29] to address the ASTE task.
- **IMN-IOG.** Similar to Peng-unified-R + LOG, in [19], the IMN [30] and IOG [29] are combined together to generate the final triplets.

End-to-end methods

It	has	a	bad	memory	but	a	great	battery	life	
N	N	N	N	N	N	N	N	N	N	It
	N	N	N	N	N	N	N	N	N	has
		N	N	N	N	N	N	N	N	a
			O	Neg	N	N	N	N	N	bad
				A	N	N	N	N	N	memory
					N	N	N	N	N	but
						N	N	N	N	a
							O	Pos	Pos	great
								A	A	battery
									A	life

Fig. 3. A tagging example with GTS. This sentence includes two triplets, ($memory, bad, negative$) and ($battery\ life, great, positive$) respectively.

Table 1
Detailed information in the datasets released in [19]. “#S”, “#A”, “#O”, and “#T” indicate the number of sentences, aspect terms, opinion terms, and triplets, respectively.

Datasets	14res			14lap			15res			16res		
	Train	Dev	Test	Train	Dev	Test	Train	Dev	Test	Train	Dev	Test
#S	1259	315	493	899	225	332	603	151	325	863	216	328
#A	2064	487	851	1257	332	467	871	205	436	1213	298	456
#O	2098	506	866	1270	313	478	966	226	469	1329	331	485
#T	2356	580	1008	1452	383	547	1038	239	493	1421	348	525

Table 2
Results on the dataset ASTE-Data-V2 released in [18]. The best results under the F1 metric are in bold. †: The results are retrieved from ([18]). §: The results are retrieved from [32].

Methods	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
CMLA + †	39.18	47.13	42.97	30.09	36.92	33.16	34.56	39.84	37.01	41.34	42.10	41.72
RINANTE + †	31.42	39.38	34.95	21.72	18.66	20.07	29.88	30.06	29.97	25.68	22.30	23.87
Li-unified-R†	41.04	67.35	51.00	40.56	44.28	42.34	44.72	51.39	47.82	37.33	54.51	44.31
Peng-unified-R + PD†	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§	63.00	55.10	58.70	49.20	40.50	45.10	57.90	42.70	48.90	60.30	53.40	56.50
JET(M = 6)†	61.50	55.13	58.14	53.03	33.89	41.35	64.37	44.33	52.50	70.94	57.00	63.21
PASTE-AF§	62.40	61.80	62.10	53.70	48.60	51.00	54.80	53.40	54.10	62.20	62.80	62.50
PASTE-OF§	63.40	61.90	62.60	59.70	48.10	50.00	54.80	52.60	53.70	62.30	63.60	62.90
DGEIAN_V2	71.68	61.62	**66.26**	60.15	43.44	**51.14**	61.84	50.99	**55.89**	69.40	60.15	**64.37**

Table 3
Model comparison results on the datasets released in [19]. The best results under the F1 metric are in bold. *: The results are retrieved from ([19]).

Methods	14res			14lap			15res			16res		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Peng-unified-R + LOG*	58.89	60.41	59.64	48.62	45.52	47.02	51.70	46.04	48.71	59.25	58.09	58.67
IMN + IOG*	59.57	63.88	61.65	49.21	46.23	47.68	55.24	52.33	53.75	-	-	-
GTS-CNN*	70.79	61.71	65.95	55.93	47.52	51.38	60.09	53.57	56.64	62.63	66.98	64.73
GTS-BiLSTM*	67.28	61.91	64.49	59.42	45.13	51.30	63.26	50.71	56.29	66.07	65.05	65.56
DGEIAN	71.03	62.63	**66.55**	60.74	45.56	**51.72**	64.87	52.75	**57.11**	69.07	65.64	**67.30**

- **OTE-MTL.** A framework [31] based on multi-task learning is proposed to obtain the triplet elements, where the sentiment of each triplet is parsed with a bi-affine scorer.
- **JET.** In [18], a brand-new approach named JET derived from the novel position-aware tagging scheme is proposed to predict the triplets for the ASTE task.
- **GTS.** A unified tagging scheme [19] is demonstrated to deal with triplet extraction. It also introduces a fresh inference strategy according to the mutual indications among the triplet elements.
- **PASTE.** The authors of [32] propose a tagging-free solution built on an encoder-decoder architecture to produce all triplets, where the decoding framework is derived from the pointer network.

3.3. Experiments Settings

Similar to the recent research [19], we unite the 300-dimension domain-general embedding from GloVe [33] and the domain-specific embedding trained with fastText [34] to obtain the initial word embeddings. Note that the 300-dimension domain-general embedding is trained with 840 billion tokens. The initial values of POS embeddings are obtained from a random sampling of the standard normal distribution, and they are learnable in the training process. The dropout rate of these two embeddings is 0.3 and the dimension of POS embeddings is 100. The hidden state size of the Bi-LSTM and GCN is 300. We adopt Adam optimizer [35] to optimize networks with an initial learning rate of 0.001. The dropout rate in the interactive attention mechanism is 0.1 to avoid over-fitting. We use Spacy[2] to generate dependency graphs. The POS of each word of the sentence is tagged by the NLTK[1]. The batch size is 32 and the set of development is applied to the early stop. We determine the best model parameters in terms of the metric of F1 of the development set and report the average of 5 runs of random initialization. We release our code in https://github.com/sll0107/DGEIAN.

3.4. Experimental Results

Performance comparisons in metrics of P, R, and F1 on the datasets released in [18,19] are respectively reported in Table 2 and Table 3, which crystallize a conclusion that our model significantly outperforms all other alternatives on all datasets under the metric of F1. Besides, the performance of our model exceeds other approaches in both the metrics of P and R in most cases. As shown in Table 2, compared with the best pipeline method Peng-unified-R + PD, our model achieves obvious improvement on F1 scores. Compared with the tagging-free scheme method PASTE, our model outperforms by 3.66, 0.14, 1.79, and 1.47 F1 points on four datasets, which is a strong demonstration of the superiority of our presented DGEIAN. Note that we select the better result from the two strategies of PASTE as the comparison. The observations in Table 3 represent that our DGEIAN performs better than the strong pipeline models Peng-unified-R + IOG and IMN + LOG on the four datasets. Especially, our method outperforms GTS-BiLSTM in all metrics on the four datasets and acquires 2.06, 0.42, 0.82, and 1.74 improvements in the F1 on four datasets, which strongly proves the effectiveness of grammatical relationships. These comparisons illustrate that our approach can benefit from the implicit interaction between the triplet elements.

3.5. Ablation Study

We conduct an ablation study to further examine the effectiveness of different modules of DGEIAN. The performance under the metric of F1 is reported in Table 4. "POS" indicates that we remove the POS embeddings, and"IA" denotes that we remove the interaction attention mechanism and feed its initial input to the attention layer. It can be observed that an average drop of 3.7 in F1 scores of the four datasets when our model did not establish an interactive attention mechanism. We can thus conclude that the interactive process is critical for our model. When POS embedding is not used, an average drop of 1.21 in F1 scores of the four datasets can be observed. And these experimental results prove their effectiveness for the ASTE task.

Table 4
Overall ablation results under the metric of F1 on four datasets released in [19].

Ablation	14res	14lap	15res	16res
DGEIAN	66.55	51.72	57.11	67.30
-POS	65.51	49.38	56.69	66.28
-IA	62.41	47.44	54.11	63.92

3.6. Effect of Combinations of POS

Motivated by the existing research conclusions, it is reasonable to assume that the combination of nouns, adjectives, verbs, and adverbs vitally affects triplet exaction. In order to clarify the role

Table 5
Results of DGEIAN and its variants under the metric of F1 on four datasets released in [19]. For simplicity, nouns, adjectives, adverbs, and verbs in this table are expressed in abbreviations.

DGEIAN Variants	Combination of POS	14res	14lap	15res	16res
DGEIAN-I	(n, adj, others)	64.83	48.47	55.93	65.92
DGEIAN-II	(n, adj, v, others)	64.14	47.51	55.35	65.03
DGEIAN-III	(n, adj, adv, others)	64.96	48.84	56.10	65.80
DGEIAN	(n, adj, v, adv, others)	66.55	51.72	57.11	67.30

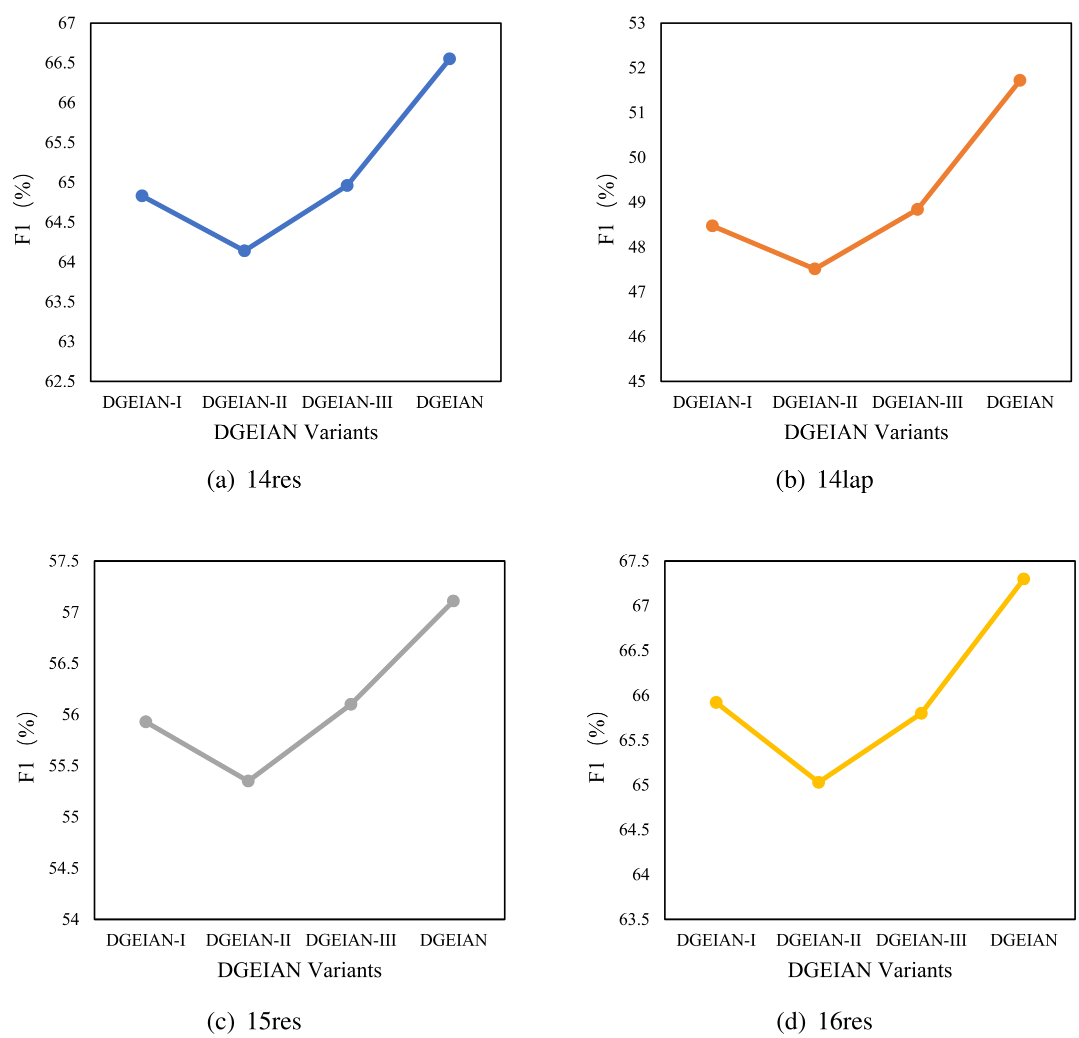

Fig. 4. F1 score comparison of DGEIAN and its variants on four datasets released in [19].

of different combinations of POS categories in triplet extraction, we design the other three DGEIAN variants and train them on the four datasets. Table 5 exhibits the performances of these variants. The corresponding F1 scores carves are plotted in Fig. 4. As shown in Fig. 4, it can be confirmed that the DGEIAN achieves better model performance compared with the other DGEIAN variants. Hence, we can conclude that the combination of nouns, adjectives, verbs, and adverbs provides the largest assistance for triplet exaction. In addition, we notice that when the verb category is added to DGEIAN-I, the prediction accuracy has declined, while the performance reaches the peak when both verb and adverb categories are combined with DGEIAN-I simultaneously. This phenomenon can be attributed to the fact that adverbs are often used to modify verbs in sentence structure, considering only verbs will seriously affect the performance of parsing and lead to performance degradation.

3.7. Impact of Iteration Number

To figure out a proper iteration number of the interactive attention mechanism, we conduct the study and the performance is reported in Fig. 5. We can observe that three is the best iteration number for all datasets and more iterations are detrimental to the effectiveness of our model. Hence, we can further conclude that

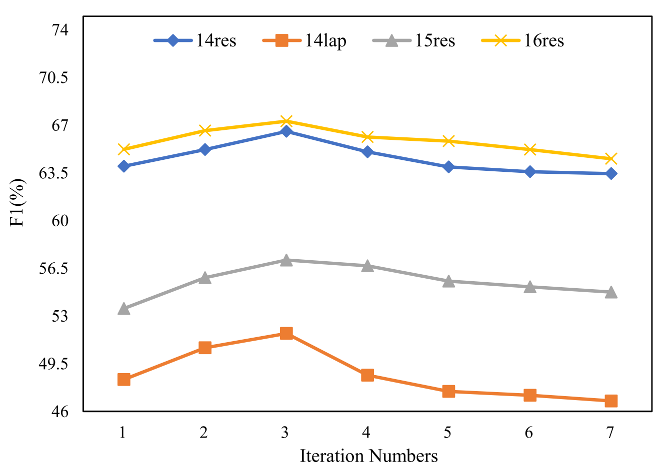

Fig. 5. A demonstration of F1- Iteration Numbers curves on four datasets released in [19] respectively.

a too-small number of iterations would lead to the syntactic dependence information not being fully integrated by contextual representation, while a too large number of iterations would make the model unable to learn the most representative features.

3.8. Case Study

To illustrate the differences between GTS and our model, we provide a case study in Table 6, including several predictable examples, the corresponding ground truth, and prediction results of these two models. The first sample contains two triplets and the first triplet has a multi-word term. GTS cannot correctly recognize the boundary of the multi-word aspect term *staff members* when predicting the first triplet, nor can it predict the second triplet. Similarly, the second example includes two triplets, and the first triplet consists of a multi-word term. GTS fails to extract the triplet with the multi-word aspect term *Indian food*. The third sample is a simple case with complicated word order. GTS mistakenly predicts *crime* as an opinion term. The fourth sample includes three triplets and the distance between any two triplets is very close. GTS fails to accurately predict the correspondence between the aspect term and the opinion term. However, our model can accurately handle the above examples, which further proves that DGEIAN is capable of achieving the proper balance between syntactic dependencies and semantic information. In particular, in the fifth example, although both our model and GTS mistakenly predict the span of aspect term, our model successfully predicts the corresponding sentiment polarity due to introducing the abundant grammatical relationships.

4. Related Work

ASTE focuses on constructing a complete solution for ABSA, aiming to extract all aspect terms, the corresponding opinion terms, and the sentiment polarity of each aspect term from a comment. Once the task was put forward, it received wide attention. The existing work can be summarized as follows.

ASTE is first investigated in [1], which employs a two-stage pipeline approach. However, this pipeline method is easily adversely affected by error propagation. In order to avoid the disadvantages of pipeline methods, the subsequent work [19] constructs a new grid tagging scheme to extract triplets

Table 6
Case study.

Example	Golden Truth	GTS	DGEIAN
From the beginning, we were met by friendly staff members, and the convenient parking at Chelsea Piers made it easy for us to get to the boat.	(staff members, friendly, positive)	(staff, friendly, positive)	(staff members, friendly, positive)
	(parking, convenient, positive)		(parking, convenient, positive)
If you 're craving some serious Indian food and desire a cozy ambiance, this is quiet and exquisite choice.	(Indian food, serious, positive)	(ambiance, cozy, positive)	(Indian food, serious, positive)
	(ambiance, cozy, positive)		(ambiance, cozy, positive)
Two wasted steaks – what a crime!	(steaks, wasted, negative)	(steaks, wasted, negative) (steaks, crime, positive)	(steaks, wasted, negative)
Anytime and every time I find myself in the neighborhood I will go to Sushi Rose for fresh sushi and great portions all at a reasonable price.		(sushi, fresh, positive)	
	(sushi, fresh, positive) (portions, great, positive) (price, reasonable, positive)	(sushi, great, positive) (portions, fresh, positive) (portions, great, positive) (price, reasonable, positive)	(sushi, fresh, positive) (portions, great, positive) (price, reasonable, positive)
One caveat: Some of the curried casseroles can be a trifle harsh.	(curried casseroles, harsh, negative)	(curried casseroles, trifle harsh, positive)	(curried casseroles, trifle harsh, negative)

simultaneously. Similarly, another novel position-aware tagging scheme [18] is proposed to extract triplets in an end-to-end manner. Due to ASTE can be divided into three subtasks, the multi-task learning framework is also used to solve ASTE. The authors of [31] achieve a multi-task learning framework and apply heuristic rules to generate the final triplets. The authors of [36] construct a novel multi-task learning framework and creatively apply multiple tagging mechanisms in the process of subtasks learning. Different from the methods based on tagging scheme, an end-to-end tagging-free solution [32] is introduced to tackle the shortage of previous tagging-based methods. Recently, the authors of [37] propose a novel solution, which divides the relationship between two words in a sentence into ten types, and introduces language features into the modeling process. Some works address ASTE at the span-level. A novel span-sharing joint extraction method [38] extracts all the spans in a sentence to generate all possible pairs of aspect terms and opinion terms, and then obtains the corresponding sentiment according to each candidate pair and its context. Another span-level model [39] distinguishes the direction from aspect term to opinion term and from opinion term to aspect term and a novel span separation loss is proposed. Moreover, many modeling paradigms have also been employed for tackling ASTE. The authors of [40] formulate ASTE as a typical machine reading comprehension problem and solve it in the pipeline. The authors of [41] take a similar approach by formalizing the ASTE task as a multi-turn machine reading comprehension task. In [42], the authors formulate ASTE into a unified index generation problem and build it on the pre-training sequence-to-sequence model. The authors of [43] convert ASTE into a text generation task and design two generation paradigms. The authors of [44] transform ASTE into an unordered triplet set recognition problem and build it on the encoder-decoder architecture. Nevertheless, most of the existing methods do not pay attention to the rich grammatical information of sentences. Motivated by this, we concentrate on the abundant grammatical relationships between words to capture the associations among triplet elements.

5. Conclusions

We put forward an innovative approach DGEIAN for ASTE. Different from the existing works, we take advantage of multiple grammatical relationships, including syntactic dependencies, contextual features, and POS information. For greater use of the syntactic dependencies and contextual features among triplet elements, we design an interactive attention mechanism jointly considering the contextual representations and syntactic representations in an iterative interaction manner. Moreover, we explicitly model the information of some specific POS categories to enhance semantic features. Our approach significantly outperforms the previous methods for ASTE and our analysis demonstrates the effectiveness of the sub-module of DGEIAN.

Since this work does not distinguish dependency types when obtaining syntactic information, the dependency types will be considered in future work. Moreover, the domain-specific knowledge will also be investigated to further enhance the performance of our DGEIAN. Next, extending our DGEIAN to support other subtasks of ABSA is also challenging yet significant.

CRediT authorship contribution statement

Lingling Shi: Conceptualization, Methodology, Software. **Donghong Han:** Data curation, Writing - original draft. **Jiayi Han:** Visualization, Investigation. **Baiyou Qiao:** Software, Validation. **Gang Wu:** Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (61672144), the National Key R&D Program of China (2019YFB1405302) and the Fundamental Research Funds for the Central Universities under Grant(N2016009).

References

[1] H. Peng, L. Xu, L. Bing, F. Huang, W. Lu, L. Si, Knowing what, how and why: A near complete solution for aspect-based sentiment analysis, in: The Thirty-Fourth AAAI Conference on Artificial Intelligence, 2020, pp. 8600–8607. doi:10.1609/aaai.v34i05.6383.

[2] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, Semeval-2014 task 4: Aspect based sentiment analysis, in: Proceedings of the 8th International Workshop on Semantic Evaluation, 2014, pp. 27–35, https://doi.org/10.3115/v1/s14-2004.

[3] Y. Yin, F. Wei, L. Dong, K. Xu, M. Zhang, M. Zhou, Unsupervised word and dependency path embeddings for aspect term extraction, in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, 2016, pp. 2979–2985, https://doi.org/10.48550/arXiv.1605.07843.

[4] X. Li, L. Bing, P. Li, W. Lam, Z. Yang, Aspect term extraction with history attention and selective transformation, in: Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018, pp. 4194–4200, https://doi.org/10.24963/ijcai.2018/583.

[5] D. Ma, S. Li, F. Wu, X. Xie, H. Wang, Exploring sequence-to-sequence learning in aspect term extraction, in: Proceedings of the 57th Conference of the Association for Computational Linguistics, 2019, pp. 3538–3547, https://doi.org/10.18653/v1/p19-1344.

[6] B. Yang, C. Cardie, Extracting opinion expressions with semi-markov conditional random fields, in: Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 2012, pp. 1335–1345.

[7] R. Klinger, P. Cimiano, in: Joint and pipeline probabilistic models for fine-grained sentiment analysis: Extracting aspects, subjective phrases and their relations, in 13th IEEE International Conference on Data Mining Workshops, 2013, pp. 937–944, https://doi.org/10.1109/ICDMW.2013.13.

[8] B. Yang, C. Cardie, Joint inference for fine-grained opinion extraction, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, 2013, pp. 1640–1649.

[9] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, K. Xu, Adaptive recursive neural network for target-dependent twitter sentiment classification, in: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 2014, pp. 49–54, https://doi.org/10.3115/v1/p14-2009.

[10] M. Zhang, Y. Zhang, D. Vo, Gated neural networks for targeted sentiment analysis, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016, pp. 3087–3093.

[11] M. Yang, W. Tu, J. Wang, F. Xu, X. Chen, Gated neural networks for targeted sentiment analysis, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017, pp. 5013–5014.

[12] X. Li, L. Bing, W. Lam, B. Shi, Transformation networks for target-oriented sentiment classification, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018, pp. 946–956, https://doi.org/10.18653/v1/P18-1087.

[13] J. Tang, Z. Lu, J. Su, Y. Ge, L. Song, L. Sun, J. Luo, Progressive self-supervised attention learning for aspect-level sentiment analysis, in: Proceedings of the 57th Conference of the Association for Computational Linguistics, 2019, pp. 557–566, https://doi.org/10.18653/v1/p19-1053.

[14] X. Li, L. Bing, P. Li, W. Lam, A unified model for opinion target extraction and target sentiment prediction, in: The Thirty-Third AAAI Conference on Artificial Intelligence, 2019, pp. 6714–6721. doi:10.1609/aaai.v33i01.33016714.

[15] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, 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, 2019, pp. 504–515, https://doi.org/10.18653/v1/p19-1048.

[16] Z. Chen, T. Qian, Relation-aware collaborative learning for unified aspect-based sentiment analysis, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 3685–3694, https://doi.org/10.18653/v1/2020.acl-main.340.

[17] H. Zhao, L. Huang, R. Zhang, Q. Lu, H. Xue, 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, 2020, pp. 3239–3248, https://doi.org/10.18653/v1/2020.acl-main.296.

[18] L. Xu, H. Li, W. Lu, L. Bing, Position-aware tagging for aspect sentiment triplet extraction, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 2339–2349, https://doi.org/10.18653/v1/2020.emnlp-main.183.
[19] Z. Wu, C. Ying, F. Zhao, Z. Fan, X. Dai, R. Xia, Grid tagging scheme for aspect-oriented fine-grained opinion extraction, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 2576–2585, https://doi.org/10.18653/v1/2020.findings-emnlp.234.
[20] K. Shuang, M. Gu, R. Li, J. Loo, S. Su, Interactive pos-aware network for aspect-level sentiment classification, Neurocomputing 420 (2021) 181–196, https://doi.org/10.1016/j.neucom.2020.08.013.
[21] W. Khong, L. Soon, H. Goh, S. Haw, Leveraging part-of-speech tagging for sentiment analysis in short texts and regular texts, in: Semantic Technology - 8th Joint International Conference, Vol. 11341, 2018, pp. 182–197. doi:10.1007/978-3-030-04284-4_13.
[22] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735.
[23] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in: 5th International Conference on Learning Representations, 2017, https://doi.org/10.48550/arXiv.1609.02907.
[24] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, I. Androutsopoulos, Semeval-2015 task 12: Aspect based sentiment analysis, in: Proceedings of the 9th International Workshop on Semantic Evaluation, 2015, pp. 486–495, https://doi.org/10.18653/v1/s15-2082.
[25] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. Al-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O.D. Clercq, V. Hoste, M. Apidianaki, X. Tannier, N.V. Loukachevitch, E.V. Kotelnikov, N. Bel, S.M.J. Zafra, G. Eryigit, Semeval-2016 task 5: Aspect based sentiment analysis, in: Proceedings of the 10th International Workshop on Semantic Evaluation, 2016, pp. 19–30, https://doi.org/10.18653/v1/s16-1002.
[26] W. Wang, S.J. Pan, D. Dahlmeier, X. Xiao, Coupled multi-layer attentions for co-extraction of aspect and opinion terms, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, 2017, pp. 3316–3322.
[27] H. Dai, Y. Song, Neural aspect and opinion term extraction with mined rules as weak supervision, in: Proceedings of the 57th Conference of the Association for Computational Linguistics, 2019, pp. 5268–5277, https://doi.org/10.18653/v1/p19-1520.
[28] X. Li, L. Bing, P. Li, W. Lam, A unified model for opinion target extraction and target sentiment prediction, in: The Thirty-Third AAAI Conference on Artificial Intelligence, 2019, pp. 6714–6721. doi:10.1609/aaai.v33i01.33016714.
[29] Z. Fan, Z. Wu, X. Dai, S. Huang, J. Chen, 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, 2019, pp. 2509–2518, https://doi.org/10.18653/v1/n19-1259.
[30] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, 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, 2019, pp. 504–515, https://doi.org/10.18653/v1/p19-1048.
[31] C. Zhang, Q. Li, D. Song, B. Wang, A multi-task learning framework for opinion triplet extraction, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, 2020, pp. 819–828, https://doi.org/10.18653/v1/2020.findings-emnlp.72.
[32] R. Mukherjee, T. Nayak, Y. Butala, S. Bhattacharya, P. Goyal, PASTE: A tagging-free decoding framework using pointer networks for aspect sentiment triplet extraction, in: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, 2021, pp. 9279–9291, https://doi.org/10.18653/v1/2021.emnlp-main.731.
[33] J. Pennington, R. Socher, C.D. Manning, Glove: Global vectors for word representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014, pp. 1532–1543, https://doi.org/10.3115/v1/d14-1162.
[34] P. Bojanowski, E. Grave, A. Joulin, T. Mikolov, Enriching word vectors with subword information, Trans. Assoc. Comput, Linguistics 5 (2017) 135–146, https://doi.org/10.48550/arXiv.1607.04606.
[35] D.P. Kingma, J. Ba, Adam: A method for stochastic optimization, in: 3rd International Conference on Learning Representations, 2015, pp. 1051–1060, https://doi.org/10.48550/arXiv.1412.6980.
[36] F. Chen, Z. Yang, Y. Huang, A multi-task learning framework for end-to-end aspect sentiment triplet extraction, Neurocomputing 479 (2022) 12–21, https://doi.org/10.1016/j.neucom.2022.01.021.
[37] H. Chen, Z. Zhai, F. Feng, R. Li, X. Wang, Enhanced multi-channel graph convolutional network for aspect sentiment triplet extraction, in: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, 2022, pp. 2974–2985.
[38] Y. Li, Y. Lin, Y. Lin, L. Chang, H. Zhang, A span-sharing joint extraction framework for harvesting aspect sentiment triplets, Knowl. Based Syst. 242 (2022), https://doi.org/10.1016/j.knosys.2022.108366 108366.
[39] Y. Chen, K. Chen, X. Sun, Z. Zhang, Span-level bidirectional cross-attention framework for aspect sentiment triplet extraction, CoRR abs/2204.12674. doi:10.48550/arXiv.2204.12674.
[40] Y. Mao, Y. Shen, C. Yu, L. Cai, A joint training dual-mrc framework for aspect based sentiment analysis, in: Thirty-Fifth AAAI Conference on Artificial Intelligence, 2021, pp. 13543–13551. doi:10.48550/arXiv.2101.00816.
[41] S. Chen, Y. Wang, J. Liu, Y. Wang, Bidirectional machine reading comprehension for aspect sentiment triplet extraction, in: Thirty-Fifth AAAI Conference on Artificial Intelligence, 2021, pp. 12666–12674. doi:10.48550/arXiv.2103.07665.
[42] H. Yan, J. Dai, T. Ji, X. Qiu, Z. Zhang, 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, 2021, pp. 2416–2429, https://doi.org/10.18653/v1/2021.acl-long.188.
[43] W. Zhang, X. Li, Y. Deng, L. Bing, W. Lam, 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, 2021, pp. 504–510, https://doi.org/10.18653/v1/2021.acl-short.64.
[44] H. Fei, Y. Ren, Y. Zhang, D. Ji, Nonautoregressive encoder-decoder neural framework for end-to-end aspect-based sentiment triplet extraction, IEEE Transactions on Neural Networks and Learning Systems (2021) 1–13, https://doi.org/10.1109/TNNLS.2021.3129483.

Lingling Shi received her B.S. degree from the Department of Information Management and Information System, Liaoning University in 2020. She is currently working towards her Master degree in Department of Computer Science and Technology at Northeastern University. Her research interests include deep learning, text mining and sentiment analysis.

Donghong Han received the Ph.D. degree from Northeastern University, China in 2007. She is currently a professor of the School of Computer Science and Engineering, Northeastern University, China. Her current research interests include data flow management, uncertain data flow analysis and social network sentiment analysis. She is a member of China Computer Federation (CCF), and a member of Chinese Information Processing Society of China (CIPSC).

Jiayi Han received the B.S. degree from Northeastern University, China in 2018. He is currently pursuing a Ph. D. degree in the Institute of Science and Technology for Brain-Inspired Intelligence at Fudan University, China. His research interests include facial expression recognition (FER) and 3D processing. His has contributed papers on AAAI and Neurocomputing.

Baiyou Qiao received the Ph. D degree in Computer Software and Theory from Northeastern University, China in 2006. Since 2008, he has been an Associate Professor in Northeastern University, China. He is a member of China Computer Federation (CCF). His main research interests include cloud computing, Big Data mining and analysis, spatial temporal data management and etc.

Gang Wu received his BS and MS degrees from Northeastern University, China in 2000 and 2003, respectively, and his PhD degreefrom Tsinghua University, China in 2008. He is an associate professor of the School of Computer Science and Engineering at Northeastern University, China. His main research interests include main memory database, knowledge graph, and social networks. He is a member of ACM, a member of Chinese Information Processing Society of China (CIPSC), and a member of China Computer Federation (CCF).