Neurocomputing 492 (2022) 211–221

Contents lists available at ScienceDirect

Neurocomputing

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

Span-based dual-decoder framework for aspect sentiment triplet extraction

Yuqi Chen [a,b,c,d], Zequn Zhang [a,b], Guangyao Zhou [a,b,*], Xian Sun [a,b,c], Keming Chen [a,b]

[a] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China
[b] Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China
[c] University of Chinese Academy of Sciences, Beijing100190, China
[d] School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 7 July 2021
Revised 14 February 2022
Accepted 3 April 2022
Available online 6 April 2022
Communicated by Zidong Wang

2020 MSC:
00-01
99-00

Keywords:
Aspect sentiment triplet extraction
ASTE
Span-based

ABSTRACT

Aspect sentiment triplet extraction (ASTE) aims to extract aspects from review sentences along with their corresponding opinions and sentiments to form opinion triplets. Since each factor within a sentiment triplet could be a single word or a phrase, defining and implementing the span-level features for this span-level task is critical and challenging. However, prior works typically formulate the ASTE task as a sequence tagging problem and address it with token-level models, limiting the extraction performances of long entities and suffering from cascading errors due to sequential decoding. Although some methods have enumerated all possible spans as input, they fail to explicitly build the interaction among the potential triplets and semantic information within the sentence explicitly. To address these problems, we propose a span-based joint training framework, where each potential entity is represented as an independent span and sentiment polarity is classified by using the corresponding independent span representations. Specifically, we design two different transformer-based decoders to extract the aspects and their corresponding opinions, respectively. Those decoders utilize multiple multi-head attention mechanisms to model the associations among the spans and the semantic information between the spans and the sentences. To verify the effectiveness of our approach, we conduct extensive experiments on four benchmark datasets. The experimental results demonstrate that our proposed method significantly outperforms existing state-of-art methods.

1. Introduction

Aspect based sentiment analysis (ABSA) is essential in natural language processing (NLP) which analyzes people's detailed insights towards a product or service. ABSA has three fundamental subtasks: aspect term extraction (ATE)[1–6], opinion term extraction (OTE)[7–9], and aspect-level sentiment classification (ASC) [10,11]. Existing works generally solve these subtasks individually or combine them as aspect polarity co-extraction task (APCE) [12,13] and aspect-opinion pair extraction task (AOPE)[14,15]. More recently, [16] presents the aspect sentiment triplet extraction (ASTE) task, which aims to identify aspect terms from review sentences along with their corresponding opinions and sentiments. As shown in Fig. 1, ASTE task aims to identify triplets such as (*hard drive, quiet, positive*), where "*hard drive*" is the aspect term, "*quiet*" represents the corresponding opinion term, and "*positive*" is the sentiment polarity.

Some recent research [16–22] formulated the ASTE task as a sequence tagging problem and solved it with sequential methods, which already achieved a great process. However, there still exist some limitations. Firstly, for an aspect term (or opinion term) consisting of multiple words, existing sequential methods predict a label for each of these words separately, which is challenging to use the global information of the aspect term (or opinion term) and suffer from cascading errors. For instance, there is a chance that the sequential method regard "*hard*" as a meaning of "*difficult*" and regard "*drive*" as a verb due to ignoring the global semantic information. Secondly, since the entities may contain multiple words, the sentiment correspondence may not be guaranteed [13]. For example, the aspect term "*boots up*" should be predicted with the same and correct label "*Positive*" to obtain the precisely correct aspect tag prediction. In general, these methods are

* Corresponding author at: Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China.

E-mail address: chenyuqi19@mails.ucas.ac.cn (G. Zhou).

https://doi.org/10.1016/j.neucom.2022.04.022
0925-2312/

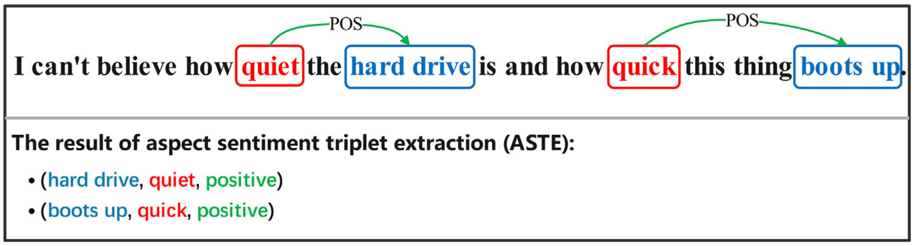


Fig. 1. An example of the ASTE task. The spans highlighted in blue are aspects. The spans in red are opinions. Sentiment polarities are marked with green.

sequential in nature and have difficulty directly modelling the span-based interactions among the aspect and opinion spans.

Another alternative approach for the ASTE task is to use the span-based model. [23] proposes a model that generates explicit representation and independently predicts the sentiment relation for all possible aspects and opinions spans. Although this method has shown encouraging performance, it has ignored the interaction among the potential triplets and the semantic information within the sentence. Furthermore, different triplets are mutually dependent and indicative within the sentence. For instance, if we already know "*quiet*" is an opinion term, then "*hard drive*" is more appropriate as an aspect term to form the triplet, while "*boots up*" is discarded because it does not match "*quiet*" in the current context.

In this paper, we propose a span-based dual decoder framework to extract sentiment triplets jointly. In detail, there are two parts in our framework: a span representation generator and a dual-decoder framework. First, the proposed framework utilizes the BERT[24] model as the encoder to learn token-level representations in the sentence and generates the representations for all possible spans. Since the generator models each possible span independently and parallel, our model does not have cascading errors. Then, to build the association among the potential triplets and retrieve the semantic information in the sentence, both the aspect decoder and the opinion decoder utilize two multi-head attention mechanisms to model the relation among the spans and exploit the potential bridges between the spans and whole sentence. Besides, the opinion decoder contains an extra multi-head attention mechanism to restrict the predicted opinion terms only pair-wise with a specific aspect term. In the training process, our model predicts the sentiment twice in both two decoders instead of predicting sentiment polarity only once as in prior research[16,18,23], and we design an algorithm to choose a better one during inference. To validate the effectiveness of our model, we conduct a series of experiments based on public datasets. The comparison results show that the proposed method outperforms the state-of-the-art methods in ASTE, APCE, and AOPE tasks.

In summary, the main contributions of this paper are concluded as follows:

(1) This paper proposes a practical span-based dual-decoder framework for ASTE task, which can take advantage of span-level information to overcome the drawbacks of sequential decoding. Through the dual-decoder module, our framework can extract aspect sentiment triplet in a joint training framework.
(2) We utilize multiple multi-head attention mechanisms to form the dual-decoder framework. By this careful design, our model can effectively build associations among potential triplets, fuse semantic information from the sentence, and transfer information between two decoders.
(3) Our proposed method yields state-of-the-art results on two benchmark datasets, and we perform various experiments to verify the effectiveness of the method.

2. Related Work

2.1. Aspect based sentiment analysis

Aspect based sentiment analysis (ABSA) is a long-standing natural language process task that consists of various tasks, including aspect term extraction (ATE) [25,26,5,4,7,2], opinion term extraction (OTE) [27,28,4,9], aspect-level sentiment classification (SC) [29–32]. Some researchers pay more attention to subtasks with compound output, such as aspect-oriented opinion extraction (AOE)[8,9], aspect term polarity co-extraction (APCE)[33–35,13,**?**], and aspect–opinion pair extraction (AOPE)[14,15,36]. However, the above compound tasks are still not enough to get a complete picture regarding sentiment. To alleviate this issue, [16] presents the aspect sentiment triplet extraction (ASTE) task, which aims to identify aspects, opinions, and sentiments in a unified framework and proposes a two-stage framework to identify those three factors. However, this method is inefficient as it has two stages and three separate models. To deal with this problem, [17,18] propose two different tagging schemes for the ASTE task. [17] specifies the structural information for a triplet by tagging the relative position of opinions and the sentiment in the beginning position of aspect tags. [18] proposes a grid tagging scheme that tags the relations between all word-pairs and addresses the ASTE task in an end-to-end way. [19] utilized machine reading comprehension to identify triplets in two directions. [21] formalizes the ASTE task as a paraphrase generation problem that exploits the semantics of the sentiment elements and generates the triplets in the natural language form. [22] proposes a hierarchical reinforcement learning framework to focus on sentiments expressed in a sentence first and then identify the aspect and the opinion terms for these sentiments. However, these methods suffer from several disadvantages because of the sequence tagging scheme. Our model eliminates the burden of sentiment label inconsistency and inadequate entity semantics issues by utilizing span-level information.

2.2. Span-based Approaches

Recently, the span-based methods have achieved highly competitive performance in NLP tasks. Some works [37,38] use span representations derived from a BiLSTM over concatenated ELMo, word and character embeddings towards joint entity and relation extraction. While [38] focuses on joint entity and relation extraction, [37] conducts a beam search over the hypothesis space, estimating which spans participate in entity classes, relations and coreferences. Moreover, [39] proposes an end-to-end model to address coreference resolution by considering all spans within a document as potential mentions. [15] proposes a multi-task learning framework to solve the AOPE task (also named PAOTE) based on shared spans, where the terms are extracted under the supervision of span boundaries, and the pair-wise relations are jointly identified using the span representations. The closest work to ours is [23]. Similar to [15], the fundamental idea of their approaches is to use span-level representations to extract the aspect terms and opinion terms and identify the pair-wise relations between every potential aspect-opinion pair with the span representations. Although the generation of all possible span representations of our approach is similar to [23], the decoding of our approach is quite different from their method, and we will detail the differences in Section 3.2.2.

3. Method

In this section, we first define the ASTE task in Section 3.1. Then detail our approach in Section 3.2, which contains a span representation generator and dual-decoder framework. Finally, we introduce the inference process in Section 3.3.

3.1. Task Definition

Given a sentence $X = \{x_1, x_2, \cdots, x_n\}$ with n tokens as the input, the ASTE task aims to extract all annotated triplets $T = \{(a_k, o_k, p_k)\}_{k=1}^{|T|}$ of the given sentence X, where $a, o, p, |T|$ refer to the aspect, the opinion, the sentiment polarity, and the number of triplets, respectively. Note that the aspects and the opinions could be a single word or a phrase. Beyond that, the number of aspects may not equal the opinions because an aspect may pair up with multiple opinions, vice versa. Therefore, instead of considering the ASTE task as a sequence labelling problem, it should be formalized as a joint span classification problem. Specifically, we enumerate all the possible spans $S = \{s_1, s_2, \cdots, s_m\}$ from the given sentence X, where each span is a slice of up to length L. Based on the candidate spans, the ASTE task can be decomposed into two subtasks which are defined as follows:

- **Aspect Polarity Co-Extraction Task**
  Let $\mathscr{T} = \{POS, NEU, NEG\}$ denote a set of pre-defined sentiment types. The aspect polarity co-extraction task is, for each span $s_i \in S$, to predict an entity type as an aspect label $y_{as}(s_i) \in \mathscr{T}$ or $y_{as}(s_i) = null$ which denoting span s_i is not an aspect. The output of the aspect polarity co-extraction task is $Y_{as} = \{(s_i, p_i) : s_i \in S, p_i \in \mathscr{T}\}$.
- **Aspect-oriented Opinion Polarity Co-Extraction Task**
  Let S_{as} denote the set of aspect spans. The aspect-oriented opinion polarity co-extraction task is, for every given aspect span $s_j \in S_{as}$ and every possible span $s_i \in S$, to predict an entity type $y_{op}(s_j, s_i) \in \mathscr{T}$ as an opinion label or the span s_i is not an opinion. The output of the task is $Y_{op} = \{(s_j, s_i, p_j) : s_j \in S_{as}, s_i \in S, p_j \in \mathscr{T}\}$.

3.2. Framework

The overall architecture of our span-based dual-decoder framework is shown in Fig. 2. Our model consists of a span generator and a dual-decoder module which consists of an aspect decoder and an opinion decoder. Given an input sentence, the span generator enumerates all possible spans of the sentence based on the outputs of the BERT[24] encoder. For a joint training setup, the span representations are shared for two decoders. The aspect decoder takes the span representations to predict the aspects and the corresponding sentiments (i.e. aspect polarity co-extraction). Meanwhile, the opinion decoder uses all the span representations generated by BERT and the span representations of each aspect term as the input, then predicts the opinion terms of each aspect term and classifies their corresponding sentiments (i.e. aspect-oriented opinion polarity co-extraction).

3.2.1. Span Representation Generator

Given an input sentence $\{x_1, x_2, \cdots, x_n\}$, the BERT encoder is adopted to generate the hidden representations sequence $\mathbf{H}$:

$$\mathbf{H} = \{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_n\} \tag{1}$$

A span $s_i = \{x_{START(i)}, \cdots, x_{END(i)}\}$ represents a single word or a phrase with a starting index $START(i)$ and an ending index $END(i)$. The span vector representation s_i is generated by averaging over all hidden representations of the token within the span:

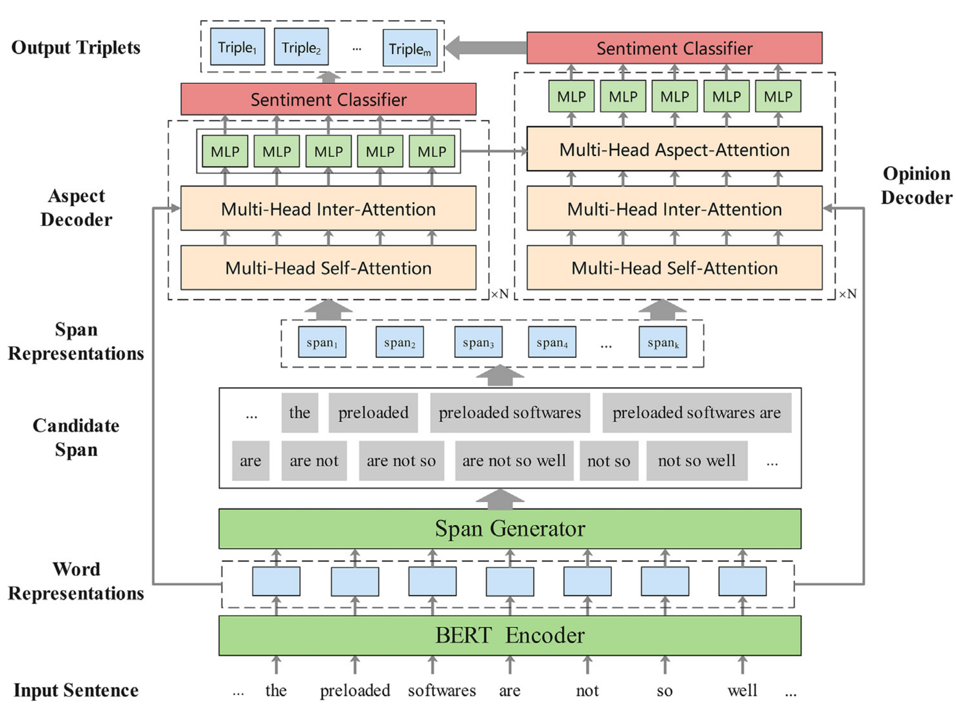


Fig. 2. The architecture of our proposed Span-based Dual-decoder Framework.

$$s_i = meanpooling(\mathbf{h}_{START(i)}, \mathbf{h}_{START(i)+1}, \cdots, \mathbf{h}_{END(i)}) \tag{2}$$

where the maximum length of each span s_i should not exceed L:

$$\begin{aligned} &1 \leqslant START(i) \leqslant END(i) \leqslant n \\ &0 \leqslant END(i) - START(i) \leqslant L \end{aligned} \tag{3}$$

The whole span representation $\mathbf{P}_S$ is:

$$\mathbf{P}_S = \{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_m\} \tag{4}$$

where m is the number of the candidate spans.

3.2.2. Dual-decoder Framework

In order to effectively use the contextualized information of the span representation $\mathbf{P}_S$, we introduce a dual-decoder framework to extract triples in two cascade steps. Firstly, an aspect decoder is adopted to find the aspect terms in all possible spans S and determine the sentiments of corresponding aspect terms. Then for each aspect found, we apply an opinion decoder to extract all opinions of the given aspect and determine the sentiments for corresponding opinions. Both two decoders are composed of a stack of N identical basic decoder blocks, respectively. In each block, there are several multi-head attention mechanisms. The details of the two decoders are given in the following sections.

Aspect Decoder

The aspect decoder aims to predict all possible aspects in the input sentence X by decoding the span representation s_i. More precisely, all span representation s_i is fed into the aspect decoder that consists of several blocks. In each block, a multi-head self-attention mechanism is employed to model the relationship between potential aspect spans, then a multi-head inter attention mechanism is utilized to fuse the information of the given sentence into the candidate spans. For the multi-head self-attention mechanism in the k-th stacked decoder block, the result is calculated as:

$$u^{as,self}_{i,k} = FFNN^{as,self}_k(\hat{s}_i, \theta^k_{as,self}) \tag{5}$$

$$\alpha^{as,self}_{i,k} = \frac{\exp(\hat{s}_i)}{\sum_{\hat{s}_j \in \mathbf{P}_S} \exp(\hat{s}_j)} \tag{6}$$

$$s^{as,self}_{i,k} = \sum_{\hat{s}_j \in \hat{\mathbf{P}}_S} \alpha^{as,self}_{i,k} \cdot \hat{s}_j \tag{7}$$

where $\hat{s}_i$ is the input in the k-th stacked decoder block, $\hat{s}_i = s_i$ when $k = 1$, and $\hat{s}_i = s^{as,self}_{i,k-1}$ when $k \neq 1$. And $\theta^k_{as,self}$ represents the parameter for the feed forward neural network (FFNN) in the multi-head self attention mechanism of the k-th stacked decoder block.

After that, we employ a multi-head inter attention mechanism to build the association among the span representations and the whole sentence. The multi-head inter attention mechanism uses the result of multi-head self-attention mechanism $\mathbf{P}^{as}_{self}$ and BERT hidden representations of all tokens $\mathbf{H}$ in the sentence as the input:

$$u^{as,inter}_{i,k} = FFNN^{as,inter}_k(s^{as,self}_i, \theta^k_{as,inter}) \tag{8}$$

$$\alpha^{as,inter}_{i,k} = \frac{\exp(u^{as,inter}_{i,k})}{\sum_{\mathbf{h}_j \in \mathbf{H}} \exp(\mathbf{h}_j)} \tag{9}$$

$$s^{as,inter}_{i,k} = u^{as,inter}_{i,k} + \sum_{\mathbf{h}_j \in \mathbf{H}} \alpha^{as,inter}_{i,k} \cdot \mathbf{h}_j \tag{10}$$

where $\theta_{as,inter}$ is the parameter for the feed forward neural network (FFNN) in the multi-head inter attention mechanism.

After the calculations of the last block is completed, each span representation $s^{as,inter}_i$ of the output is fed into an FFNN that work with a non-linear activation, and then normalized with the softmax function to output the aspect sentiment label score of each span with its confidence score by:

$$q^{as}_i = FFNN^{as}_{\mathcal{T}}\left(s^{as,inter}_i, \theta_{as}\right) \tag{11}$$

$$P_{as}(c|s_i) = Softmax\left(q^{as}_i\right) \tag{12}$$

where θ_{as} is the parameter for the $FFNN^{as}_{\mathcal{T}}$ and $c \in \{POS, NEU, NEG, null\}$.

Opinion Decoder

The opinion decoder aims to extract all opinions and classify the sentiment polarity from the input sentence X for a specific aspect. Similar to the aspect decoder, the opinion decoder takes all candidate span representations s_i and the span representation of a specific aspect term s^{as} as the input. In addition, compared to the aspect decoder, each block of the opinion decoder has an additional multi-head attention mechanism to model the association between the candidate span and the given aspect term. Specifically, the multi-head self-attention mechanism in the k-th stacked opinion decoder block is calculated as follows:

$$u^{op,self}_{i,k} = FFNN^{op,self}_k(\check{s}_i, \theta^k_{op,self}) \tag{13}$$

$$\alpha^{op,self}_{i,k} = \frac{\exp(\check{s}_i)}{\sum_{\check{s}_j \in \mathbf{P}_S} \exp(\check{s}_j)} \tag{14}$$

$$s^{op,self}_{i,k} = \sum_{\check{s}_j \in \mathbf{P}_S} \alpha^{op,self}_{i,k} \cdot \check{s}_j \tag{15}$$

where $\check{s}_i$ is the input in the k-th stacked opinion decoder block, $\check{s}_i = s_i$ when $k = 1$, and $\check{s}_i = s^{op,self}_{i,k-1}$ when $k \neq 1$. And $\theta^k_{op,self}$ represent the parameter for the feed forward neural network (FFNN) in the multi-head self attention mechanism of the k-th stacked opinion decoder block.

Next, similar to the aspect decoder, we adopt a multi-head inter attention mechanism to model the relation between these potential opinion spans and the whole sentence:

$$u^{op,inter}_{i,k} = FFNN^{op,inter}_k(s^{op,self}_i, \theta^k_{op,inter}) \tag{16}$$

$$\alpha^{op,inter}_{i,k} = \frac{\exp(u^{op,inter}_{i,k})}{\sum_{\mathbf{h}_j \in \mathbf{H}} \exp(\mathbf{h}_j)} \tag{17}$$

$$s^{op,inter}_{i,k} = u^{op,inter}_{i,k} + \sum_{\mathbf{h}_j \in \mathbf{H}} \alpha^{op,inter}_{i,k} \cdot \mathbf{h}_j \tag{18}$$

where $\theta_{op,inter}$ is the parameter for the feed forward neural network (FFNN) in the multi-head inter attention mechanism of the k-th stacked opinion decoder block.

To restrict the extraction results of the opinion term in opinion decoder to correspond only to one given aspect term, we further employ an additional multi-head restrict attention mechanism over the span representations and the span representation of a specific aspect term s^{as}:

$$u^{op,res}_{i,k} = FFNN^{op,res}_k(s^{op,inter}_{i,k}, \theta^k_{op,res}) \tag{19}$$

$$\alpha^{op,res}_{i,k} = \frac{\exp(u^{op,inter}_{i,k})}{\exp(s^{as})} \tag{20}$$

$$s^{op,res}_{i,k} = u^{op,res}_{i,k} + \alpha^{op,res}_{i,k} \cdot s^{as} \tag{21}$$

where $\theta_{op,res}$ is the parameter for the feed forward neural network (FFNN) in the multi-head restrict attention mechanism of the k-th stacked opinion decoder block.

After the calculations of the last block of the opinion decoder, the span representation $s_i^{op,res}$ is fed into the feed-forward network and then normalized with the softmax function to output the unnormalized score as well as the probability of the term label:

$$q_i^{op} = FFNN_{\mathcal{T}}^{op}(s_i^{op,res}, \theta_{op}) \tag{22}$$

$$P_{op}(c|s_i) = Softmax(q_i^{op}) \tag{23}$$

where θ_{op} is the parameter for the $FFNN_{\mathcal{T}}^{op}$ and $c \in \{POS, NEU, NEG, null\}$.

Training

For aspect decoder, the loss function can be defined using the span-level cross entropy cost function between the predicted distribution $P_{as}(c|s_i)$ and the gold distribution $P_{as}(c^*|s_i)$:

$$\mathcal{L}_{as} = -\sum_{i}^{m} P_{as}(c^*|s_i) log(P_{as}(c|s_i)) \tag{24}$$

Then, for the opinion decoder, the loss function can be defined by using the cross-entropy cost function:

$$\mathcal{L}_{op} = -\sum_{i}^{m}\sum_{j}^{\mathbf{P}_s^{as}} P_{op}\left(c^*|s_i, s_j^{as}\right)\left(P_{op}\left(c|s_i, s_j^{as}\right)\right) \tag{25}$$

where $P_{op}(c^*|s_i)$ is the gold distribution of a opinion term given the respect to a specific aspect term and $P_{op}(c|s_i)$ is the predicted result of the corresponding opinion term, $\mathbf{P}_s^{as}$ denote the set of the ground truth of the aspect terms. Eventually, losses from the aspect decoder and the opinion decoder are combined as the joint training objective of the entire framework:

$$\mathcal{L} = \mathcal{L}_{op} + \mathcal{L}_{as} \tag{26}$$

Differences from Span-ASTE Our approach differs from the Span-ASTE approach in the following ways: (1) We use the dual-decoder framework to extract the aspect terms and all the opinion terms for a specific aspect term. In this case, our framework does not have to classify all the aspect-opinion pairs. (2) We adopt several attention mechanisms to explicitly model the relationships among the spans and the semantic connections between spans and the whole sentence. In addition, there are also attention mechanisms that restrict the opinion decoder only to decode opinion terms for a specific aspect term. (3) we classify the sentiments for both the aspect terms and the opinion terms and design a simple algorithm to select a better sentiment polarity based on the confidence score. As we would show in the experiments, our model achieves a competitive performance in all the benchmark datasets, using the identical pre-trained encoders (BERT).

3.3. Inference Process

Our framework has several differences between the training process and the inference process. During the training process, the ground truth of all aspects is already known. Thus, the training process uses the correct aspects for training. However, without the golden truth during the inference process, our model infers the triplets in a pipeline as shown in Algorithm 1. The aspects A and their corresponding aspect sentiments C^{as} are predicted by the aspect decoder. After that, the opinions O and their corresponding opinion sentiments C^{op} are predicted by the opinion decoder with respect to the aspects A. Since our model classifies the sentiment in both aspect decoder and opinion decoder, the final triplets T are formed after judging the triplet's sentiment c_j by the scores of each aspect sentiment p_i^{as} and opinion sentiment p_j^{op}.

Algorithm 1:The Inference Process for Aspect Sentiment Triplets Extraction task of the Span-based Dual Decoder Framework

Input: sentence X
Output: triplets $T = \{(a, o, s)\}$
  Initialize $T = \{\}$;
  Use X to generate all possible span representations s_i described in Section 3.2.1;
  Input all possible spans s_i to aspect decoder, output the aspect span candidates a_i, the corresponding aspect sentiments c_i^{as} and the score of each aspect sentiment q_i^{as};
  If $A = \{\}$, return T;
for $a_i \in A$**do**
    Input all span representations s_i and the span representations s_i^{as} that there corresponding sentiments are in $\mathcal{T} = \{POS, NEU, NEG\}$ to opinion docoder, output the opinion spans o_j, the opinion sentiment c_j^{op} and the score of each opinion sentiment q_j^{op};
    if $q_i^{as} < q_j^{op}$ **then**
      triplet sentiment $c_j = c_j^{op}$
    else
      triplet sentiment $c_j = c_i^{as}$
    end if
    $T \leftarrow T \cup \{(a_i, o_j, c_j), j = 1, 2, \cdots\}$
  end for

4. Experiments

4.1. Datasets

In order to study aspect-opinion pair extraction, [8] releases four aspect-opinion pair datasets based on SemEval Challenges [40–42]. However, their datasets are not suitable for ASTE task due to the lack of the sentiment polarity annotation. The original SemEval Challenge only provides the annotation of aspects and the corresponding sentiment. Thus [16] combine the datasets released by [8] and original SemEval Challenge datasets to form ASTE datasets, called ASTE-DATA-V2 [1]. The detailed statistics for the ASTE-DATA-V2 datasets are shown in Table 1.

4.2. Experimental Settings

We use the BERT-base model for the encoding layer with 12 attention heads, 12 hidden layers, and the hidden size of 768, resulting in 110 M pretrained parameters. The maximum length of generated spans is set to 8. During training, we adopt Adam optimizer[43] for optimization. The learning rate for training decoders and the fine-tuning rate for BERT is set to 1e-4 and 1e-5, respectively. We set the batch size to 8 and the dropout rate to 0.1. We select the best model parameters based on the best F1 score on the development data and apply it to the test data for evaluation. Note that all the baseline methods are implemented using publicly released source codes. All the compared models are trained with the best settings, and the results for test sets are reported when it achieves the best performances on the dev sets. We run our model on three GeForce RTX 2080ti GPUs and train our model for 120 epochs in about three hours.

[1] https://github.com/xuuuluuu/SemEval-Triplet-data/tree/master/ASTE-Data-V2-EMNLP2020

Table 1
Statistics of ASTE-DATA-V2 datasets annotated by [16]. The "LAP" and "RES" represent restaurant and laptop domain datasets.

Datasets		Sentence	Positive	Neutral	Negative	Single-word	Muti-word
14LAP	Train	1266	1692	166	480	1586	752
	Dev	310	404	54	119	388	189
	Test	492	773	66	155	657	337
14RES	Train	906	817	126	517	824	636
	Dev	219	169	36	141	190	156
	Test	328	364	63	116	291	252
15RES	Train	605	783	25	205	678	335
	Dev	148	185	11	53	165	84
	Test	322	317	25	143	297	188
16RES	Train	857	1015	50	329	918	476
	Dev	210	252	11	76	216	123
	Test	326	407	29	78	344	170

4.3. Evaluation

To comprehensively measure the performances of different methods, we use *precision*, *recall*, and *F1-score* as the evaluation metrics to evaluate the results on four tasks, including aspect polarity co-extraction(APCE), aspect opinion pair extraction and aspect sentiment triplet extraction. The extracted aspects and opinions are regarded as correct only if the predicted term exactly matches a gold term.

4.4. Baselines

The following state-of-the-art (SoTA) models have been compared in the experiments.

- **Peng-two-stage** [16] is a two-stage pipeline model proposed by [16]. This model is motivated by **Li-unified-R** and fuses GCN to capture dependency information to co-extract targets with sentiment and opinion spans simultaneously at the first stage. At the second stage, it pairs up the extraction results into triplets via a relation classifier.
- **JET** [17] utilizes unified tags to solve the ASTE task as a sequence tagging problem. There are two variants of **JET**: **JETt** predict the target, the sentiment of target and the corresponding opinion. **JETo** predict the opinion, the sentiment of the opinion and the corresponding target.
- **GTS**[18] is a latest model which proposes a grid tagging schema to identify the aspect sentiment triplets in an end-to-end way. It also utilizes BERT as the encoder and designs an inference strategy to exploit mutual indication between different opinion factors.
- **Dual-MRC** [20] is a joint training model which consists of two machine reading comprehensions. One of the MRC is for aspect term extraction, and another is for aspect-oriented opinion term extraction and sentiment classification.
- **B-MRC** [19] formalizes the ASTE task as a multi-turn machine reading comprehension task, and proposes three types of queries to extract targets, opinions and the sentiment polarities of aspect-opinion pairs, respectively.
- **PARAPHRASE** [21] formalizes the ASTE task as a paraphrase generation problem. This method can exploit the semantics of the sentiment elements and generate the triplets in the natural language form.
- **ASTE-RL** [22] regards the aspect and opinion terms as arguments of the expressed sentiment in a hierarchical reinforcement learning (RL) framework which focus on sentiments expressed in a sentence first, then identify the aspect and the opinion terms for that sentiment.
- **Span-ASTE** [22] considers all possible spans in a sentence to consider the interaction between the whole spans of aspect terms and opinion terms when predicting their sentiment relation. They also propose a dual-channel span pruning strategy to ease the high computational cost caused by span enumeration.

4.5. Main Results

The experimental results for the APCE, AOPE and ASTE tasks are shown in Table 2. According to the results, our model has outperformed the best baselines by an average of 7.7% F1 score on the APCE task and an average of 2.2% F1 score on the AOPE task. Although some of the precision scores are lower than B-MRC and Dual-MRC methods, our results of recall score significantly outperform the previous baselines, which shows the higher prediction completeness of our framework than the baselines. Besides, we observe that BERT-based models perform better than the LSTM-based model. This is mainly because BERT can learn richer context semantics than LSTM. So, we only conduct the experiment of our model with BERT. B-MRC and Dual-MRC achieve better performance than JET and PENG-two-stage because they formalize the ASTE task as a multi-turn machine reading comprehension task.

Moreover, our model surpasses the best baseline in the ASTE task in both 14LAP and 14RES datasets and only achieve slightly lower results on 15RES and 16RES datasets. Compared with the sequence tagging approach, our span-based models show superior performance. This is mainly because sequence tagging methods are sequential and suffer from cascading errors. It also demonstrates that providing accurate span representation of the entities is valid, and our model can benefit from modelling the span-span and span-sentence relations. It is worth noting that our model achieves more performance gains in the ASTE task than the AOPE task compared to other models, which verifies the effectiveness of our model in the sentiment classification task.

4.6. Ablation Study

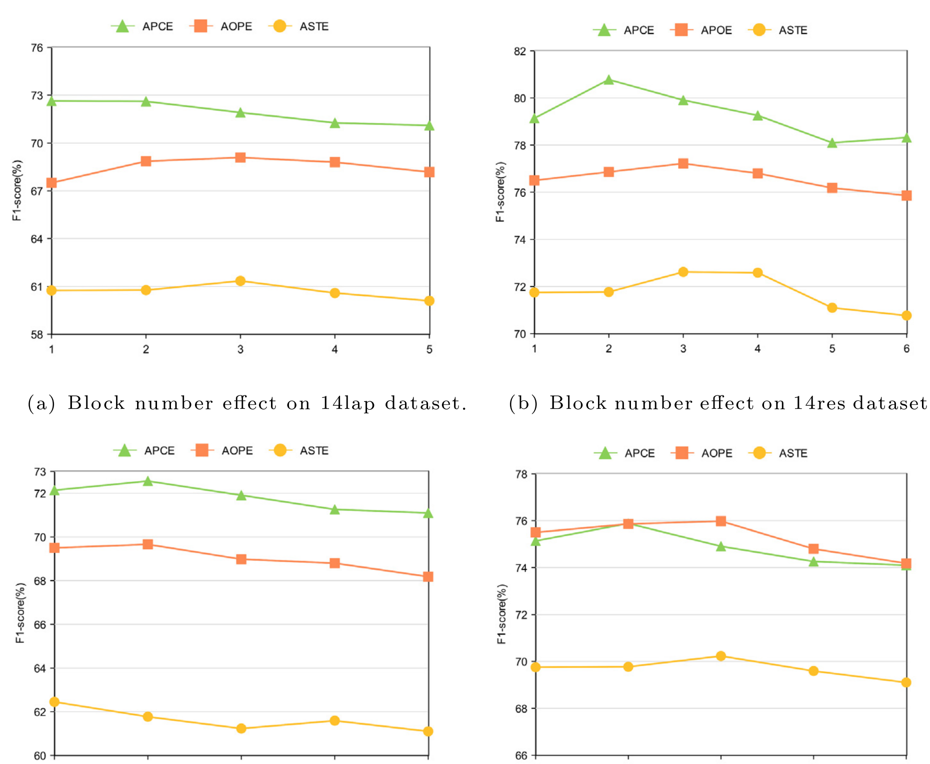
To examine the effect of the multiple attention mechanisms in the decoder blocks and sentiment selection in the inference process, we conduct several experiments with different decoder blocks and inference options. As shown in Table 3, the performance of w/o SELF ATT and w/o INTER ATT decrease significantly. The reason for this could be that both the representations of other spans and the complete integration of sentence information make essential contributions to the classification of each span. Moreover, the different results of w/o INTER ATT and w/o SELF ATT indicate that the association between spans and the sentence is more critical than other spans' information in 14LAP and 14RES datasets, but not as crucial as spans' information in 15RES and 16RES datasets. It is probably due to the domain difference and the various difficul-

Table 2
Precision (%), Recall (%) and F1 score (%) on the APCE, AOPE, and ASTE task. The best results are marked in bold. ∗ indicates that the result is reproduces by us.

Tasks	Models	14LAP			14RES			15RES			16RES		
		P	R	F1	P	R	F1	P	R	F1	P	R	F1
APCE	PENG-two-stage	63.15	61.55	62.34	74.41	73.97	74.19	67.65	64.02	65.79	71.18	72.30	71.73
	Dual-MRC	67.45	61.96	64.59	76.84	76.31	76.57	66.84	63.52	65.14	69.18	72.59	70.84
	B-MRC	**76.60***	57.18*	65.48*	81.39*	69.28*	74.85*	**77.2***	58.81*	66.76*	**77.20***	69.38*	73.08*
	Ours	72.45	**72.81**	**72.63**	**81.45**	**80.12**	**80.78**	71.75	**73.39**	**72.56**	69.76	**83.19**	**75.88**
AOPE	PENG-two-stage	50.00	58.47	53.85	47.76	68.10	56.10	49.22	65.70	56.23	52.35	70.50	60.04
	GTS	66.41	64.95	65.67	76.23	74.84	75.53	66.40	68.71	67.53	71.70	77.79	74.62
	Dual-MRC	65.43	61.43	63.37	76.23	73.67	74.93	**72.43**	58.90	64.97	77.06	74.41	75.71
	B-MRC	**82.42***	58.37*	68.34*	**80.6***	68.45*	74.03*	68.91*	61.31*	64.89*	**80.97***	73.33*	**77.41***
	Ours	71.10	**67.19**	**69.09**	74.09	**80.63**	**77.22**	70.16	**69.17**	**69.66**	73.37	**78.79**	75.98
ASTE	PENG-two-stage	40.40	47.24	43.50	44.18	62.99	51.89	40.97	54.68	46.79	46.76	62.97	53.62
	JETt	51.48	42.65	46.65	70.20	53.02	60.41	62.14	47.25	53.68	**72.12**	57.20	63.41
	JETo	55.39	47.33	51.04	67.97	60.32	63.92	58.35	51.43	54.67	64.77	61.29	62.98
	GTS-BERT	57.12	53.42	55.21	70.92	69.49	70.20	59.29	58.07	58.67	68.58	66.60	67.58
	Dual-MRC	57.39	53.88	55.58	71.55	69.14	70.32	63.78	51.87	57.21	68.60	66.24	67.40
	B-MRC	**70.89***	50.20*	58.78*	**75.41***	64.04*	69.26*	**69.83***	56.04*	58.74*	69.03*	66.02*	67.49*
	PARAPHRASE	-	-	61.13	-	-	72.03	-	-	62.56	-	-	**71.70**
	ASTE-RL	64.80	54.99	59.50	70.60	68.65	69.61	65.45	60.29	62.72	67.21	69.69	68.41
	Span-ASTE	63.44	55.84	59.38	72.89	70.89	71.85	62.18	**64.45**	**63.27**	69.45	71.17	70.26
	Ours	62.36	**60.37**	**61.35**	72.12	**73.14**	**72.62**	64.27	60.73	62.45	68.74	**71.79**	70.23

Table 3
Precision (%), Recall (%) and F1 score (%) on the aspect–opinion pair extraction (AOPE) task. The best results are in bold. w/o SELF ATT, w/o INTER ATT, and w/o Both represent models without self attention mechanisms, without inter attention mechanisms, and without both of them in the blocks, respectively. And S, S_{AT}, and S_{OT} means classify sentiment with confidence score, with the sentiment of aspect term, and with the sentiment of opinion term, respectively.

Method	14LAP			14RES			15RES			16RES		
	S	S_{AT}	S_{OT}	S	S_{AT}	S_{OT}	S	S_{AT}	S_{OT}	S	S_{AT}	S_{OT}
Ours	**61.35**	60.37	61.02	72.62	71.47	**72.71**	**62.45**	62.45	62.32	**70.23**	70.04	69.83
w/o SELF ATT	59.56	59.08	**59.68**	**71.31**	71.03	71.31	**59.28**	59.02	59.25	69.52	69.39	**69.54**
w/o INTER ATT	**60.37**	59.70	59.65	**71.92**	71.41	71.57	**58.32**	58.12	57.98	**68.88**	68.88	68.64
w/o Both	60.56	60.23	**60.62**	**71.86**	71.86	71.86	58.77	**58.81**	58.42	69.37	**69.39**	69.33



(a) Block number effect on 14lap dataset. (b) Block number effect on 14res dataset.

(c) Block number effect on 15res dataset. (d) Block number effect on 16res dataset.

Fig. 3. Effects of block number for APCE, AOPE, ASTE tasks.

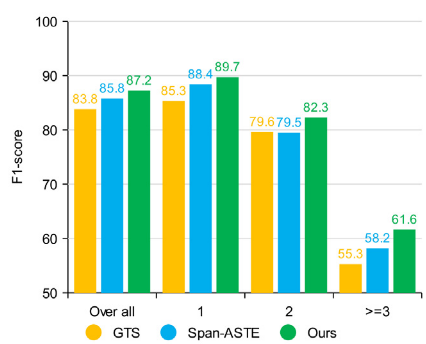
ties among different datasets. The performance of w/o Both indicates that the FFNN can also perform the function of the attention modules to some extent, but not as well as explicitly modelling the association as the attention modules do. Although sentiment selection shows inferior performance in the 14RES dataset, its experimental results perform better than the results in which sentiments were selected directly using the aspect decoder and the opinion decoder on 14LAP, 15RES and 16RES datasets. This indicates the validity of the sentiment selection S can take advantage of the sentiments of both the aspect decoder and the opinion decoder to determine better sentiments for triplets.

4.7. Effects of Block Number

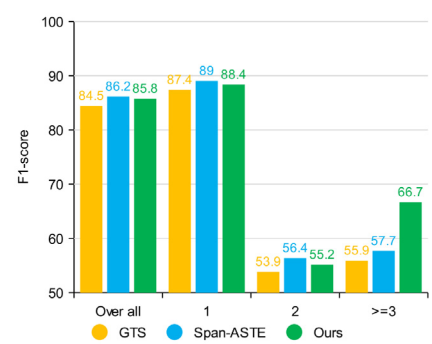
In order to investigate the importance of the decoder, we conduct the experiments by changing the number of decoder blocks and reporting the results for APCE, AOPE, ASTE tasks on the 14lap and 14res datasets. According to Fig. 3, our model achieves the best F1 score in both AOPE and ASTE tasks when the block numbers of the decoders are set to 1,1,2,3 on 14LAP, 14RES, 15RES, and 16RES, respectively. Meanwhile, our model achieves the best F1 score in the APCE task by the block number of 3,2,2,2 on 14LAP, 14RES, 15RES, and 16RES. All tasks achieve the best score of no more than three decoder blocks. We conjecture that this is probably because the excessive number of blocks of the decoders increases the number of parameters and tends to overfit the model. However, there are also advantages to blocks. With the deepening of the decoder blocks, more multi-head self-attention modules allow for better modelling of the connections between spans, more multi-head inter attention modules allow for more complete integration of sentence information into the spans, and more multi-head aspect attention modules allow for more substantial aspect restriction of opinion term extraction. Besides, the variation tendency for APCE task different from AOPE and ASTE tasks, since the APCE task does not include opinion term extraction (OTE) task, which is closely related to the number of decoder stacks, and AOPE and ASTE tasks differ only in whether or not to make a judgment of sentiment polarity which is implemented after the decoders.

4.8. Effects of Target Length

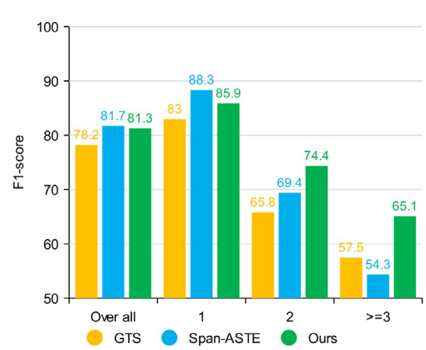
To investigate the effects of target length on the extraction task, we report the F1 scores of GTS, Span-ASTE, and our model for the extraction task in different lengths of aspects and opinions on the 14RES and 15RES datasets. The results are shown in Fig. 4. Our overall extraction performance shows that our model significantly outperforms the GTS model in both datasets and achieves competitive results compared to Span-ASTE. Although the performance of our model is slightly lower in opinion extraction of 14RES and aspect extraction of 15RES, as the target length increases, the performance gap between our framework and other models becomes more conspicuous. Since long targets are generally found in long sentences, and Span-ASTE produces more candidate aspect terms and opinion terms in long sentences, many redundant pairs swamp the correct pairs in the exhaustive pairing of candidate terms, which adversely affects the results. However, our model extracts aspect words first and then identify opinion words, where the classification of the spans is independent of each



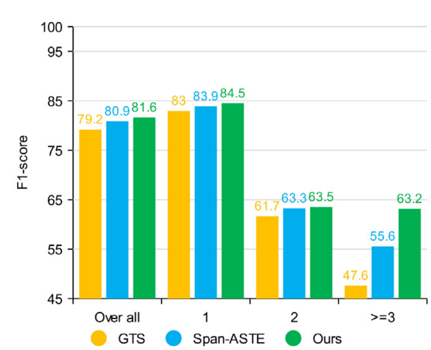
(a) Target length effect on 14RES dataset for aspect term extraction.



(b) Target length effect on 14RES dataset for opinion term extraction.



(c) Target length effect on 15RES dataset for aspect term extraction.



(d) Target length effect on 15RES dataset for opinion term extraction.

Fig. 4. Effects of target length for aspect term extraction and opinion term extraction.

Table 4
Case study of aspect sentiment triplet extraction task. Right and Wrong predictions are marked with ✓and × respectively.

Example	I have to say they have one of the fastest delivery times in the city.	It 's traditional, simple italian food.	MS Office 2011 for Mac is wonderful, well worth it.
Ground Truth	(delivery times, fastest, POS)	(italian food, traditional, POS), (italian food, simple, POS)	(MS Office 2011 for Mac, wonderful, POS), (MS Office 2011 for Mac, well worth, POS)
Our model	(delivery times, fastest, POS) ✓	(italian food, traditional, POS) ✓, (italian food, simple, POS) ✓	(MS Office 2011 for Mac, wonderful, POS) ✓, (MS Office 2011 for Mac, worth, POS) ×
GTS	(delivery, fastest, POS) ×	(italian food, simple, POS) ✓	(MS Office 2011, wonderful, POS) ×, (MS Office 2011, well worth, POS) ×

other. The decoder's self-attention and inter attention mechanisms can also help the model differentiate the spans of different lengths.

4.9. Case Study

To further validate the effectiveness of our model, we compare our model with the state-of-art model GTS based on three examples. The results are shown in Table 4. The first and the third example show that our all possible span-based method performs better in extracting entities from the sentence, especially in extracting long entities where our model has a distinct advantage. In the first example, GTS model detects "*delivery*" by mistake while our model detects "*delivery times*". In the third example, our model detects "*MS Office 2011 for Mac*" while the GTS model detects "*MS Office 2011*" by mistake. Moreover, the second example shows that our model successfully detects both "*traditional*" and "*simple*" while the GTS method only detects "*simple*", which indicates that our model can overcome the overlapping entities problems of different aspect–opinion pairs.

5. Conclusions

In this paper, we propose a span-based dual-decoder framework to handle the aspect sentiment triplet extraction (ASTE) task. Compared with previous token-level models, our framework enumerates all possible spans as input, naturally avoiding cascading errors. To decode the span representations, we deploy a dual-decoder module which consists of an aspect decoder and an opinion decoder. The multi-head attention mechanisms in both two decoders can explicitly model the relationships among the spans and exploit the potential bridges between the spans and whole sentences and guide the opinion decoder to decode only the opinion terms corresponding to specific aspect terms. We conduct extensive experiments on four widely used datasets to validate the effectiveness of our model. Experimental results show that our proposed framework outperforms the compared baselines over several ABSA tasks. This challenging task is far from being solved. We observe that both the aspect terms and the opinion terms can trigger the sentiment triplets. Our future work will concentrate on extracting the triplets with both aspect terms and opinion terms as triggers.

CRediT authorship contribution statement

Yuqi Chen: Conceptualization, Methodology, Writing - original draft, Software. **Zequn Zhang:** Formal analysis, Investigation. **Guangyao Zhou:** Visualization. **Xian Sun:** Supervision. **Keming Chen:** Data curation, Writing - review & editing, Validation, Supervision.

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.

References

[1] T.A. Rana, Y. Cheah, A two-fold rule-based model for aspect extraction, Expert Syst. Appl. 89 (2017) 273–285, https://doi.org/10.1016/j.eswa.2017.07.047.

[2] T.A. Rana, Y. Cheah, Aspect extraction in sentiment analysis: comparative analysis and survey, Artif. Intell. Rev. 46 (4) (2016) 459–483, https://doi.org/10.1007/s10462-016-9472-z.

[3] K. Liu, L. Xu, J. Zhao, Opinion target extraction using word-based translation model, in: J. Tsujii, J. Henderson, M. Pasca (Eds.), Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL 2012, July 12–14, 2012, Jeju Island, Korea, ACL, 2012, pp. 1346–1356. url:https://www.aclweb.org/anthology/D12-1123/.

[4] H. Xu, B. Liu, L. Shu, P.S. Yu, Double embeddings and cnn-based sequence labeling for aspect extraction, in: I. Gurevych, Y. Miyao (Eds.), Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15–20, 2018, Volume 2: Short Papers, Association for Computational Linguistics, 2018, pp. 592–598. doi:10.18653/v1/P18-2094. url:https://www.aclweb.org/anthology/P18-2094/.

[5] X. Li, L. Bing, P. Li, W. Lam, Z. Yang, Aspect term extraction with history attention and selective transformation, in: J. Lang (Ed.), Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13–19, 2018, Stockholm, Sweden, ijcai.org, 2018, pp. 4194–4200. doi:10.24963/ijcai.2018/583. url:https://doi.org/10.24963/ijcai.2018/583.

[6] D. Ma, S. Li, F. Wu, X. Xie, H. Wang, Exploring sequence-to-sequence learning in aspect term extraction, in: A. Korhonen, D.R. Traum, L. Màrquez (Eds.), Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers, Association for Computational Linguistics, 2019, pp. 3538–3547. doi:10.18653/v1/p19-1344. url:https://doi.org/10.18653/v1/p19-1344.

[7] T.A. Rana, Y. Cheah, Sequential patterns rule-based approach for opinion target extraction from customer reviews, J. Inf. Sci. 45 (5). doi:10.1177/0165551518808195. url:https://doi.org/10.1177/0165551518808195.

[8] Z. Fan, Z. Wu, X. Dai, S. Huang, J. Chen, Target-oriented opinion words extraction with target-fused neural sequence labeling, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2–7, 2019, Volume 1 (Long and Short Papers), Association for Computational Linguistics, 2019, pp. 2509–2518. doi:10.18653/v1/n19-1259. url:https://doi.org/10.18653/v1/n19-1259.

[9] Z. Wu, F. Zhao, X. Dai, S. Huang, J. Chen, Latent opinions transfer network for target-oriented opinion words extraction, in: The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7–12, 2020, AAAI Press, 2020, pp. 9298–9305. url:https://aaai.org/ojs/index.php/AAAI/article/view/6469.

[10] D. Ma, S. Li, X. Zhang, H. Wang, Interactive attention networks for aspect-level sentiment classification, in: C. Sierra (Ed.), Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19–25, 2017, ijcai.org, 2017, pp. 4068–4074. doi:10.24963/ijcai.2017/568. url:https://doi.org/10.24963/ijcai.2017/568.

[11] K. Sun, R. Zhang, S. Mensah, Y. Mao, X. Liu, Aspect-level sentiment analysis via convolution over dependency tree, in: K. Inui, J. Jiang, V. Ng, X. Wan (Eds.), Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing, EMNLP-IJCNLP 2019, Hong Kong, China, November 3–7, 2019, Association for Computational Linguistics, 2019, pp. 5678–5687. doi:10.18653/v1/D19-1569. url:https://doi.org/10.18653/v1/D19-1569.

[12] H. Luo, T. Li, B. Liu, J. Zhang, DOER: dual cross-shared RNN for aspect term-polarity co-extraction, in: A. Korhonen, D.R. Traum, L. Màrquez (Eds.),

Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers, Association for Computational Linguistics, 2019, pp. 591–601. doi:10.18653/v1/p19-1056. url:https://doi.org/10.18653/v1/p19-1056.

[13] 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, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019, AAAI Press, 2019, pp. 6714–6721. doi:10.1609/aaai.v33i01.33016714. url:https://doi.org/10.1609/aaai.v33i01.33016714.

[14] S. Chen, J. Liu, Y. Wang, W. Zhang, Z. Chi, Synchronous double-channel recurrent network for aspect-opinion pair extraction, in: D. Jurafsky, J. Chai, N. Schluter, J.R. Tetreault (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2020, pp. 6515–6524, https://doi.org/10.18653/v1/2020.acl-main.582.

[15] 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: D. Jurafsky, J. Chai, N. Schluter, J.R. Tetreault (Eds.), Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020, Association for Computational Linguistics, 2020, pp. 3239–3248. doi:10.18653/v1/2020.acl-main.296. url:https://doi.org/10.18653/v1/2020.acl-main.296.

[16] 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, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7–12, 2020, AAAI Press, 2020, pp. 8600–8607. url:https://aaai.org/ojs/index.php/AAAI/article/view/6383.

[17] L. Xu, H. Li, W. Lu, L. Bing, Position-aware tagging for aspect sentiment triplet extraction, in: B. Webber, T. Cohn, Y. He, Y. Liu (Eds.), Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Association for Computational Linguistics, 2020, pp. 2339–2349, https://doi.org/10.18653/v1/2020.emnlp-main.183.

[18] Z. Wu, C. Ying, F. Zhao, Z. Fan, X. Dai, R. Xia, Grid tagging scheme for aspect-oriented fine-grained opinion extraction, CoRR abs/2010.04640. arXiv:2010.04640. url:https://arxiv.org/abs/2010.04640.

[19] 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, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2–9, 2021, AAAI Press, 2021, pp. 12666–12674. url: https://ojs.aaai.org/index.php/AAAI/article/view/17500.

[20] 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, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2–9, 2021, AAAI Press, 2021, pp. 13543–13551. url:https://ojs.aaai.org/index.php/AAAI/article/view/17597.

[21] W. Zhang, Y. Deng, X. Li, Y. Yuan, L. Bing, W. Lam, Aspect sentiment quad prediction as paraphrase generation, in: M. Moens, X. Huang, L. Specia, S.W. Yih (Eds.), Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, EMNLP 2021, Virtual Event/ Punta Cana, Dominican Republic, 7–11 November, 2021, Association for Computational Linguistics, 2021, pp. 9209–9219. url:https://aclanthology.org/2021.emnlp-main.726.

[22] S.Y.B. Jian, T. Nayak, N. Majumder, S. Poria, Aspect sentiment triplet extraction using reinforcement learning, in: G. Demartini, G. Zuccon, J.S. Culpepper, Z. Huang, H. Tong (Eds.), CIKM '21: The 30th ACM International Conference on Information and Knowledge Management, Virtual Event, Queensland, Australia, November 1–5, 2021, ACM, 2021, pp. 3603–3607. doi:10.1145/3459637.3482058. url:https://doi.org/10.1145/3459637.3482058.

[23] L. Xu, Y.K. Chia, L. Bing, Learning span-level interactions for aspect sentiment triplet extraction, in: C. Zong, F. Xia, W. Li, R. Navigli (Eds.), Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, ACL/IJCNLP 2021, (Volume 1: Long Papers), Virtual Event, August 1–6, 2021, Association for Computational Linguistics, 2021, pp. 4755–4766. doi:10.18653/v1/2021.acl-long.367. url:https://doi.org/10.18653/v1/2021.acl-long.367.

[24] J. Devlin, M. Chang, K. Lee, K. Toutanova, BERT: pre-training of deep bidirectional transformers for language understanding, in: J. Burstein, C. Doran, T. Solorio (Eds.), Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2–7, 2019, Volume 1 (Long and Short Papers), Association for Computational Linguistics, 2019, pp. 4171–4186. doi:10.18653/v1/n19-1423. url:https://doi.org/10.18653/v1/n19-1423.

[25] W. Wang, S.J. Pan, D. Dahlmeier, X. Xiao, Recursive neural conditional random fields for aspect-based sentiment analysis, in: J. Su, X. Carreras, K. Duh (Eds.), Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, The Association for Computational Linguistics, 2016, pp. 616–626, https://doi.org/10.18653/v1/d16-1059.

[26] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, An unsupervised neural attention model for aspect extraction, in: R. Barzilay, M. Kan (Eds.), Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, Canada, July 30 - August 4, Volume 1: Long Papers, Association for Computational Linguistics, 2017, pp. 388–397. doi:10.18653/v1/P17-1036. url:https://doi.org/10.18653/v1/P17-1036.

[27] P. Liu, S.R. Joty, H.M. Meng, Fine-grained opinion mining with recurrent neural networks and word embeddings, in: L. Màrquez, C. Callison-Burch, J. Su, D. Pighin, Y. Marton (Eds.), Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, EMNLP 2015, The Association for Computational Linguistics, 2015, pp. 1433–1443, https://doi.org/10.18653/v1/d15-1168.

[28] S. Poria, E. Cambria, A.F. Gelbukh, Aspect extraction for opinion mining with a deep convolutional neural network, Knowl. Based Syst. 108 (2016) 42–49, https://doi.org/10.1016/j.knosys.2016.06.009.

[29] 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, ACL 2014, June 22–27, 2014, Baltimore, MD, USA, Volume 2: Short Papers, The Association for Computer Linguistics, 2014, pp. 49–54. doi:10.3115/v1/p14-2009. url:https://doi.org/10.3115/v1/p14-2009.

[30] D. Tang, B. Qin, T. Liu, Aspect level sentiment classification with deep memory network, in: J. Su, X. Carreras, K. Duh (Eds.), Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, EMNLP 2016, The Association for Computational Linguistics, 2016, pp. 214–224, https://doi.org/10.18653/v1/d16-1021.

[31] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, Exploiting document knowledge for aspect-level sentiment classification, in: I. Gurevych, Y. Miyao (Eds.), Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, July 15–20, 2018, Volume 2: Short Papers, Association for Computational Linguistics, 2018, pp. 579–585. doi:10.18653/v1/P18-2092. url:https://www.aclweb.org/anthology/P18-2092/.

[32] Z. Li, Y. Wei, Y. Zhang, X. Zhang, X. Li, Exploiting coarse-to-fine task transfer for aspect-level sentiment classification, in: The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019, AAAI Press, 2019, pp. 4253–4260. doi:10.1609/aaai.v33i01.33014253. url:https://doi.org/10.1609/aaai.v33i01.33014253.

[33] M. Mitchell, J. Aguilar, T. Wilson, B.V. Durme, Open domain targeted sentiment, in: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, EMNLP 2013, 18–21 October 2013, Grand Hyatt Seattle, Seattle, Washington, USA, A meeting of SIGDAT, a Special Interest Group of the ACL, ACL, 2013, pp. 1643–1654. url:https://www.aclweb.org/anthology/D13-1171/.

[34] 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: A. Korhonen, D.R. Traum, L. Màrquez (Eds.), Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers, Association for Computational Linguistics, 2019, pp. 504–515. doi:10.18653/v1/p19-1048. url:https://doi.org/10.18653/v1/p19-1048.

[35] H. Li, W. Lu, Learning latent sentiment scopes for entity-level sentiment analysis, in: S.P. Singh, S. Markovitch (Eds.), Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4–9, 2017, San Francisco, California, USA, AAAI Press, 2017, pp. 3482–3489. url:http://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14931.

[36] J. Zhang, Z. Zhang, Z. Guo, L. Jin, K. Liu, Q. Liu, Enhancement of target-oriented opinion words extraction with multiview-trained machine reading comprehension model, Comput. Intell. Neurosci. 6645871 (1–6645871) (2021 2021,) 13, https://doi.org/10.1155/2021/6645871.

[37] Y. Luan, L. He, M. Ostendorf, H. Hajishirzi, Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction, in: E. Riloff, D. Chiang, J. Hockenmaier, J. Tsujii (Eds.), Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2018, pp. 3219–3232, https://doi.org/10.18653/v1/d18-1360.

[38] K. Dixit, Y. Al-Onaizan, Span-level model for relation extraction, in: A. Korhonen, D.R. Traum, L. Màrquez (Eds.), Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers, Association for Computational Linguistics, 2019, pp. 5308–5314. doi:10.18653/v1/p19-1525. url:https://doi.org/10.18653/v1/p19-1525.

[39] K. Lee, L. He, M. Lewis, L. Zettlemoyer, End-to-end neural coreference resolution, in: M. Palmer, R. Hwa, S. Riedel (Eds.), Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9–11, 2017, Association for Computational Linguistics, 2017, pp. 188–197. doi:10.18653/v1/d17-1018. url:https://doi.org/10.18653/v1/d17-1018.

[40] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, Semeval-2014 task 4: Aspect based sentiment analysis, in: P. Nakov, T. Zesch (Eds.), Proceedings of the 8th International Workshop on

Semantic Evaluation, SemEval@COLING 2014, Dublin, Ireland, August 23-24, 2014, The Association for Computer Linguistics, 2014, pp. 27–35. doi:10.3115/v1/s14-2004. url:https://doi.org/10.3115/v1/s14-2004.

[41] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, I. Androutsopoulos, Semeval-2015 task 12: Aspect based sentiment analysis, in: D.M. Cer, D. Jurgens, P. Nakov, T. Zesch (Eds.), Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2015, Denver, Colorado, USA, June 4–5, 2015, The Association for Computer Linguistics, 2015, pp. 486–495. doi:10.18653/v1/s15-2082. url:https://doi.org/10.18653/v1/s15-2082.

[42] 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: S. Bethard, D.M. Cer, M. Carpuat, D. Jurgens, P. Nakov, T. Zesch (Eds.), Proceedings of the 10th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2016, San Diego, CA, USA, June 16–17, 2016, The Association for Computer Linguistics, 2016, pp. 19–30. doi:10.18653/v1/s16-1002. url:https://doi.org/10.18653/v1/s16-1002.

[43] I. Loshchilov, F. Hutter, Fixing weight decay regularization in adam, CoRR abs/1711.05101. arXiv:1711.05101. url:http://arxiv.org/abs/1711.05101.

Yuqi Chen received the bachelor's degree from Xidian University of China, Xi'an, China, in 2019. He is currently a Master candidate in the Aerospace Information Research Institute, University of Chinese Academic of Sciences, Beijing, China. His research interests include natural language processing and sentiment analysis.

Zequn Zhang received the B.Sc. degree from Peking University, Beijing, China, in 2012, and Ph.D. degree from Peking University in 2017. He is currently a Research Assistant at the Aerospace Information Innovation Institute, Chinese Academy of Science, Beijing, China. His research interests include information fusion knowledge graph and natural language processing.

Guangyao Zhou received the Ph.D degree from Chinese Academy of Science, Beijing, China. His research interests include information fusion, computer vision, and Image Change Detection.

Xian Sun received the B.Sc. degree from Beijing University of Aeronautics and Astronautics, Beijing, China, in 2004. He received the M.Sc. and Ph.D. degrees from the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, in 2009. He is currently a Professor with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. His research interests include computer vision, geospatial data mining, and remote sensing image understanding.

Keming Chen received the Ph.D degree from Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include information fusion, computer vision, and Image Change Detection.