



# Multiple graph convolutional networks for aspect-based sentiment analysis

Yuting Ma<sup>1</sup> · Rui Song<sup>2</sup> · Xue Gu<sup>3</sup> · Qiang Shen<sup>4</sup> · Hao Xu<sup>4</sup>

Accepted: 20 July 2022 / Published online: 5 October 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Aspect-based sentiment analysis is a fine-grained sentiment analysis task that identifies the sentiment polarity of different aspects in a sentence. Recently, several studies have used graph convolution networks (GCN) to obtain the relationship between aspects and context words with the dependency tree of sentences. However, errors introduced by the dependency parser and the complexity and variety of sentence structures have led to incorrect predictions of sentiment polarity. Therefore, we propose a multiple GCN (MultiGCN) model to solve this problem. The proposed MultiGCN comprises a rational GCN (RGCN) to extract syntactic structure information of sentences, a contextual encoder to extract semantic content information of sentences, a common information extraction module to combine structure and content information, and a fusion mechanism that allows interaction among the aforementioned components. Further, we propose difference and similarity losses and combine them with traditional loss function to jointly minimize the difference between the values predicted by the model and those of the labels. The experimental results show that the prediction performance of our proposed method is more than that of the state-of-the-art models.

**Keywords** Aspect-based sentiment analysis · Graph convolutional networks · Information extraction · Fusion mechanism · Loss function

## 1 Introduction

Aspect-based sentiment analysis (ABSA) [1–5] aims to determine the sentiment polarity for a specific aspect in a sentence. For example, in the sentence shown in Fig. 1, given the two aspects “fish” and “variety of fish”, the goal of ABSA is to infer the sentiment polarities for the aspect words: positive for “fish” and negative for “variety of fish”.

Many previous studies have used recurrent neural networks (RNNs) and their variants [1, 6, 7] to learn the relationships among words in a sentence. However, it cannot effectively distinguish important information in long texts. When the sentence length is long, important information may likely fade away during propagation, resulting in poor modeling. To solve this problem, the attention mechanism and its variants are widely used in this task [8–11], which can capture crucial information between aspects and context

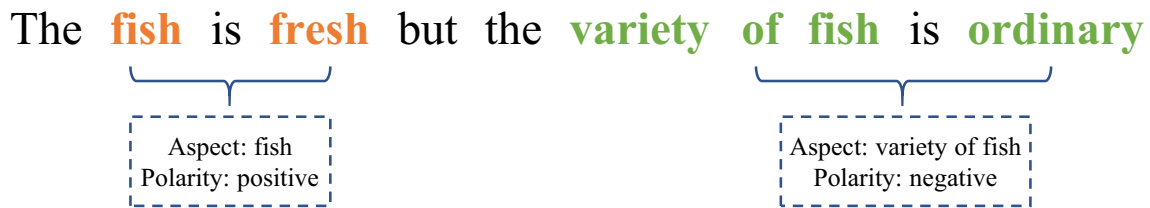
words, thereby, effectively improving the accuracy of the ABSA task. Yadav et al. [12] masked the aspect words and used Opinion Lexicon to replace the opinion words, which were trained by two Bidirectional GRUs and an attention layer. Liu et al. [13] proposed a co-attention mechanism to capture the relationships between aspect and context, where the 1-pair hop mechanism analyses the relationship between aspect and context at the lexical level and an interactive mechanism analyses the relationship between aspect and context at the feature level. Nevertheless, the risk of matching the wrong opinion words with aspect words is an unavoidable disadvantage, which leads to prediction errors. Such challenges can be alleviated to some extent by extracting local features using a convolutional neural network (CNN) [14–16]. However, this method can easily bring noise that decreases prediction accuracy as the positional relationship between the aspects and the real opinion words may be quite distant.

Based on the above problems, recent methods have introduced dependency trees to encode the structure information of sentences [17–21] and encoded the dependencies between different words using graph convolutional networks (GCN) [22–26] or graph attention networks (GAT)

---

✉ Hao Xu  
xuhao@jlu.edu.cn

Extended author information available on the last page of the article.



**Fig. 1** An example sentence with two aspects and different sentiment polarities

[27–30]. Liang et al. [31] enhanced the dependency graphs of sentences by introducing affective knowledge, and then fed the new dependency graphs into the GCN network for training. In addition, the use of GCN [32] after introducing the attention mechanism of position encoding can effectively capture sentiment dependencies of different aspects in sentences. The performance of these models goes far beyond the previous conventional models. Despite the significant improvements achieved, the ABSA task still faces great challenges. First, even the state-of-the-art parser is difficult to ensure completely correct parsing, thus the dependency tree obtained will inevitably introduce noise. Second, in the process of identifying the sentiment polarity of aspects, some sentences rely primarily on syntactic information and others on semantic information. Dealing with sentences that differ in sensitivity to syntactic and semantic information effectively is also an issue that needs to be addressed.

To overcome the aforementioned challenges, we propose a multiple GCN (MultiGCN) model that obtains the syntactic structure and the semantic content information of each sentence through RGCN and a contextual encoder, respectively. Then, the information from the RGCN and contextual encoder is combined using the common information extraction module, and the final result is obtained through a fusion mechanism. Furthermore, we modify the loss function with difference and similarity losses. The difference loss helps the model distinguish among structural information alone, content information alone, and common information after combining them. The similarity loss encourages the model to improve the degree of association between structure and content information. The experimental results confirm that they are essential for better model training.

Our contributions are summarized as follows:

- 1) We propose a MultiGCN model based on GCN, which makes full use of both structured and unstructured information of sentences and combines different types of information through a fusion mechanism.
- 2) We propose difference and similarity losses, which modify the traditional loss function and facilitate the model to learn better the degree of difference or similarity between different types of information.

- 3) We validate our model on four datasets and the results show that the MultiGCN outperforms the baseline models on all datasets.

## 2 Related work

ABSA is a direction of sentiment analysis in natural language processing [7, 33–36], which focuses sentiment on one or more specific aspects in the same sentence to judge the sentiment polarity. With regard to the structure of neural networks, existing work on ABSA can be divided into methods based on traditional deep learning and those based on graph neural networks.

### 2.1 Conventional methods

CNN [37] can extract advanced features from raw data through convolution and pooling operations. Huang et al. [14] incorporated aspect information into CNN for sentence encoding using a parameterized filter and a parameterized gate. Fan et al. [15] proposed a convolutional memory network that combines attention mechanisms and captures both words and multi-words expressions in sentences. To help the CNN feature extractor locate sentiment indicators more accurately, Li et al. [16] reviewed the drawbacks of the attention mechanism and the barriers that prevent CNN from playing a role in classification tasks, and then proposed a new classification model that uses a proximity strategy to scale the input of convolutional layers by using the positional relevance between words and aspect words.

In addition, Memory Networks [38] have been applied to this task. Tang et al. [39] applied a deep memory network to aspect level sentiment classification tasks, which uses attention mechanisms with explicit memory to capture the importance of each context word for the given aspect. Chen et al. [40] proposed a recursive attention mechanism based on memory networks to extract sentiment information separated by a long distance. The memory slices are weighted according to their position proximity to the aspect, and gated recurrent units are adopted to update the representation of aspect mentions. To better simulate sentiment interaction, Li et al. [41] integrated aspect detection into sentiment classification.

Many ABSA tasks are also modeled using RNN due to its superior capacity of sequence learning. Tang et al. [6] regarded the given aspect words as features and connected them with context features to predict sentiment polarity. Zhang et al. [1] used the gated neural network structure to model the interaction between aspect words and surrounding context words by comprehensively considering the syntax and semantics of a sentence. To further improve the accuracy of sentiment polarity discrimination, Wang et al. [8] introduced the attention mechanism and set an attention vector for each aspect based on long short-term memory (LSTM) network, which is an effective way to strengthen the neural model to focus on the relevant parts of a sentence. Subsequently, the interactive attention mechanism is explored to model the aspect-context relationship [9, 10], thus it can effectively learn important parts of sentences and aspects to provide sufficient information for judging the sentiment polarity. Fan et al. [11] combined fine-grained and coarse-grained attentions to capture aspect and context interaction at the word level, and then proposed aspect alignment loss to describe the interaction between aspects with a common context.

## 2.2 Graph neural network methods

Zhang et al. [22] constructed a GCN model on the dependency tree of sentences using syntax information and word dependences. This model starts with the BiLSTM layer and captures context information about word orders. In order to obtain aspect characteristics, GCN and masking mechanisms are used to preserve specific aspect features for predicting aspect-based sentiment polarity. In the same year, Sun et al. [23] also proposed to construct GCN on the syntactic dependency tree and combine it with BiLSTM to build the model. Later, Chen et al. [24] proposed a new gating mechanism for merging multiple tree structures in GCN coding to improve the classification accuracy of noisy texts and more effectively capture the relationship between aspect and opinion words. Furthermore, Liang et al. [25] designed an aspect-focused graph and an inter-aspect graph for each instance by considering context words related to aspect words and the dependence of aspect words on other aspects, and then conducted an ABSA task through a new interactive graph-aware model.

Recently, Zhang et al. [42] integrated word co-occurrence information and dependency type information using a hierarchical graph structure to solve the problem that most previous studies used dependency relations only but ignored different types of relations. Bai et al. [27] also used

dependency label information to distinguish the dependency types of different relations and integrated label features into the attention mechanism. They proposed a new relational GAT that improved the accuracy of parsing.

In this paper, our proposed model incorporates several different types of node features so that the sentiment polarity of a sentence can be predicted accurately. Moreover, we improve the loss function to help the model train better.

## 3 Methodology

In this section, we describe the details of our proposed model. An overview of MultiGCN is depicted in Fig. 2. For our model, a sentence  $S$  with the aspect is given, where  $S = \{w_1, w_2, \dots, a_1, \dots, a_m, \dots, w_n\}$  represents the sentence to be entered and contains the specific aspect  $\{a_1, a_2, \dots, a_m\}$ . Each word in the sentence can be found using a word embedding lookup table  $E \in R^{|V| \times d_e}$ , where  $|V|$  is the size of the word list,  $d_e$  is the dimension of word embeddings. Afterward, the resulting word embeddings are fed into a BiLSTM to encode the sentence to obtain a hidden state vector  $H = \{h_1, h_2, \dots, h_n\}$ , where  $H \in R^{2d}$  and  $d$  is the dimension of the hidden state vector of a unidirectional LSTM.

The overall architecture of our model consists of four components: the RGCN, the contextual encoder, the common information extraction module, and the fusion mechanism. First of all, the hidden representation obtained from BiLSTM is input into RGCN and contextual encoder. Then both output representations are fed into the common information extraction module to produce the new sentence representation comprehensively. Next, the fusion mechanism fuses the output of the three components, and we obtain the final feature representation by pooling and concatenation operations. Finally, the sentence  $S$  with specific aspects is predicted to have a sentiment polarity  $y \in \{1, -1, 0\}$ , where 1 is positive, -1 is negative and 0 is neutral.

### 3.1 Rational graph convolutional network

GCN can perform convolution operations on directly connected nodes in the graph structure data, which allows more global information to be obtained for each node. Therefore, by entering a dependency probability matrix, each word in a sentence can get information about the word it depends on, thus obtaining the syntactic structure information of the whole sentence. Inspired by Bai et al. [27], in order to make more comprehensive use of the results

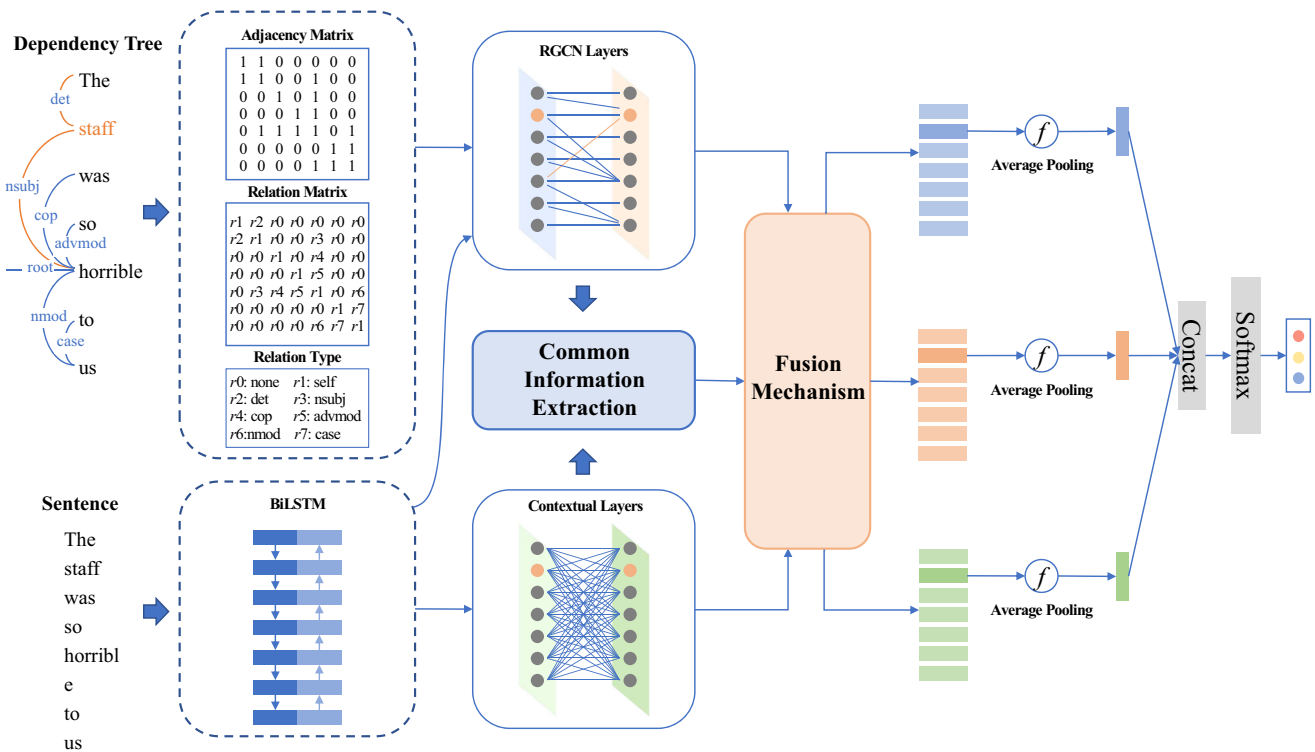


Fig. 2 The overall framework of MultiGCN

of the dependency parser, we take dependency relations and the types of relations into account and propose RGCN, as shown in Fig. 3.

To begin with, an adjacency matrix  $A_{sy} = \{a_{i,j}\}_{n \times n}$  and a relation matrix  $R = \{r_{i,j}\}_{n \times n}$  can be generated from a dependency tree, representing whether there is a dependency arc between words and what the dependency type is, respectively.  $A_{sy}$  is a 0-1 matrix, if there is an arc between two words, then  $a_{i,j} = 1$ , and  $a_{i,j} = 0$  otherwise. As for  $R$ , if  $a_{i,j} = 1$ , then  $r_{i,j} \in \{r_1, r_2, \dots, r_k\}$  is the corresponding dependency type, where  $k$  is the number of dependency types, and  $r_{i,j} = 0$  otherwise. In the next place, we convert all  $r_{i,j}$  in the relation matrix  $R$  to embedding  $e_{i,j}^r$  and combine it with the output of the BiLSTM to obtain  $H_{sy}^{(0)}$ . With the adjacency matrix  $A_{sy}$ , the structure representation  $H_{sy}$  of the sentence is expressed from the following formula:

$$H_{sy}^{(l)} = ReLU \left( A_{sy} H_{sy}^{(l-1)} W_{sy}^{(l)} \right) \quad (1)$$

where  $W_{sy}^{(l)}$  is the learnable matrix of the  $l$ -th RGCN layer.

### 3.2 Contextual encoder

As the structural information of some sentences is not obvious, the accuracy of analysis results may be reduced by only relying on the structural information extracted by

RGCN. Therefore, we consider modeling context words using another GCN, i.e., content modeling. However, unlike above, instead of using a dependency tree, we use the attention matrix  $A_{att}$  of the adjacency matrix, derived from the self-attention mechanism. The final contextual representation  $H_{con}$  of the sentence is obtained from the following formula:

$$A_{att} = \frac{H_q W_q \times H_k W_k}{\sqrt{d}} \quad (2)$$

$$H_{con}^{(l)} = ReLU \left( A_{att} H_{con}^{(l-1)} W_{con}^{(l)} \right) \quad (3)$$

where  $H_q$ ,  $H_k$ , and  $W_{con}^{(0)}$  are equal, and all of them are the output of BiLSTM;  $d$  is the dimension of the output;  $W_q$  and  $W_k$  are trainable weight matrix;  $W_{con}^{(l)}$  is the learnable matrix of the  $l$ -th GCN layer.

### 3.3 Common information extraction module

Intuitively, most sentences contain both syntactic structural information and semantic content information, which are closely related. For example, in the sentence “The staff was horrible”, there is a clear semantic relationship between “staff” and “horrible”. Meanwhile, there is also a dependent arc between the two words in the structure, so we propose a common information extraction module that combines the two features. We use the gating mechanism to integrate the

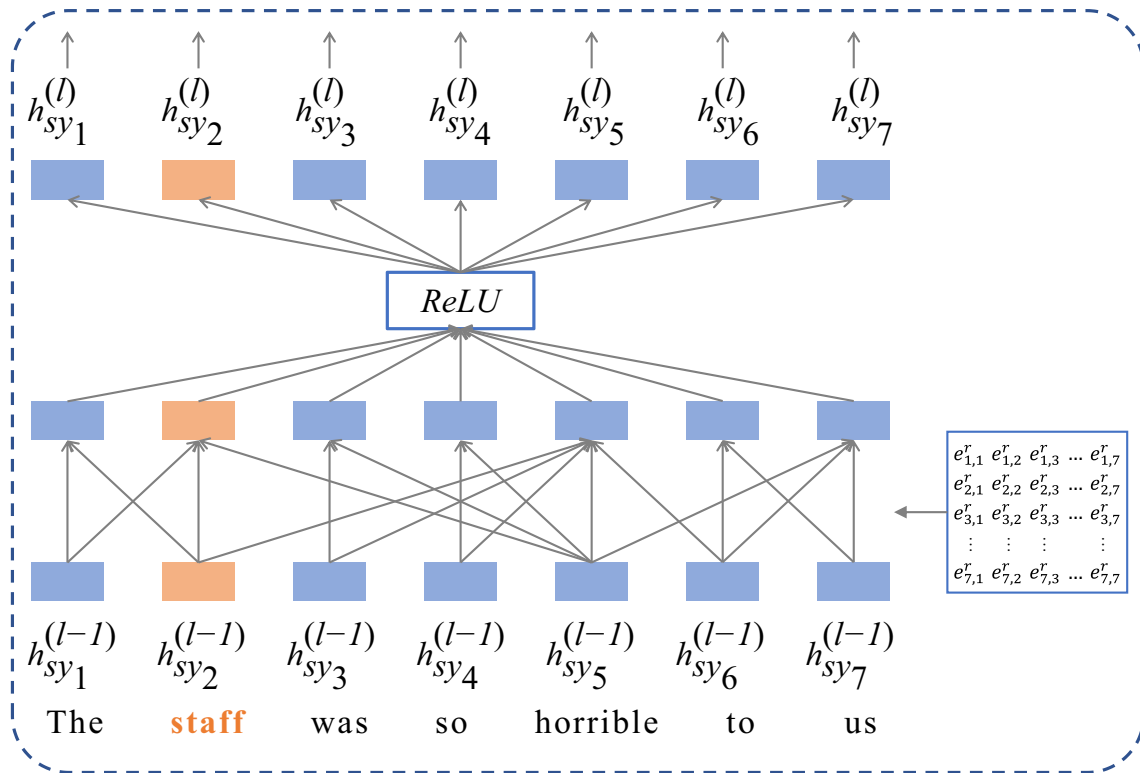


Fig. 3 Details of the RGCN structure for relation embeddings

syntactic representation  $H_{sy}$  and the content representation  $H_{con}$ , where  $H_{sy}$  is derived from (1) and  $H_{con}$  is derived from (3). Finally, the common representation  $H_c$  is obtained from the following equation:

$$H_c = g \times H_{sy} + (1 - g) \times H_{con} \tag{4}$$

$$g = \sigma ([H_{sy}, H_{con}] \times W_g + b_g) \tag{5}$$

where  $g$  is the gate;  $[H_{sy}, H_{con}]$  is the connection between  $H_{sy}$  and  $H_{con}$ ;  $W_g$  and  $b_g$  are the model weight and bias, respectively;  $\sigma$  represents the activation function sigmoid.

### 3.4 Fusion mechanism

Before obtaining the final sentence representation, we propose a fusion mechanism to extract the relevant features of different modules to obtain a more accurate representation.  $H_{sy}$ ,  $H_{con}$ , and  $H_c$  are input and interact through the following formulas:

$$H'_{sy} = softmax(H_{sy} W_1 (H_{con})^T) H_{sy} \tag{6}$$

$$H'_{con} = softmax(H_{con} W_2 (H_{sy})^T) H_{con} \tag{7}$$

$$H'_{c-sy} = softmax(H_c W_3 (H_{con})^T) H_{sy} \tag{8}$$

$$H'_{c-con} = softmax(H_c W_4 (H_{sy})^T) H_{con} \tag{9}$$

where  $W_1$ ,  $W_2$ ,  $W_3$  and  $W_4$  are trainable parameters. After that, the combination of  $H'_{c-sy}$  and  $H'_{c-con}$  gives  $H_c$  as follows:

$$H'_c = \frac{\alpha H'_{c-sy} + \beta H'_{c-con}}{2} \tag{10}$$

where  $\alpha$  and  $\beta$  are model parameters.

What's more, we use the mask mechanism and average pooling to obtain the aspect representation  $h'_{sy}$ ,  $h'_{con}$ , and  $h'_c$ , concatenating them together to obtain the final aspect representation  $h_a$ :

$$h'_{sy} = f(mask(h'_{sy1}, h'_{sy2}, \dots, h'_{syt_n})) \tag{11}$$

$$h'_{con} = f(mask(h'_{con1}, h'_{con2}, \dots, h'_{con_n})) \tag{12}$$

$$h'_c = f(mask(h'_{c1}, h'_{c2}, \dots, h'_{cn})) \tag{13}$$

$$h_a = [h'_{sy}, h'_{con}, h'_c] \tag{14}$$

where  $mask(\cdot)$  is a function of the mask mechanism, filtering the representations obtained from the different modules to get the representations about aspects;  $f(\cdot)$  is an average pooling operation.

Finally, we input the aspect representation  $h_a$  into the softmax layer to obtain the final probability distribution of sentiment polarity, completing the ABSA task:

$$y = \text{softmax}(Wh_a + b) \quad (15)$$

where  $W$  and  $b$  are the model weight and bias, respectively.

### 3.5 Model training

For optimizing model parameters to further improve model performance, we use the regularization method proposed by Li et al. [43], which is given by

$$R_O = \left\| A_{att} (A_{att})^T - I \right\|_F \quad (16)$$

$$R_D = \frac{1}{\|A_{att} - A_{sy}\|_F} \quad (17)$$

where  $I$  is an identity matrix;  $A_{att}$  and  $A_{sy}$  are the attention score matrix and adjacency matrix of the sentence, respectively.

Furthermore, we promote the traditional cross-entropy loss function by proposing the similarity and difference losses. Foremost, the similarity loss  $L_{sim}$  is proposed in that the sentence representations  $H_{sy}$  and  $H_{con}$  derived from the RGCN and contextual GCN are significantly related. Context words that are related in a sentence are also structurally connected, that is, there often have dependent arcs between the two words, so  $L_{sim}$  should be as small as possible.

$$L_{sim} = \left( H'_{con} - H'_{sy} \right)^2 \quad (18)$$

Another improvement, the difference loss  $L_{dif}$ , takes into account that the common information extraction module combines the content and structure information of a sentence to get a sentence representation  $H_c$  with two kinds of information, which should be different from the sentence representations  $H_{con}$  and  $H_{sy}$  that make use of only one kind of information. Therefore, the difference loss calculates the dot product of  $H_c$  with  $H_{con}$  and  $H_{sy}$ , respectively, and then sums them up. The smaller the value, the better the result can be obtained by making full use of the different information, as shown in the formula below:

$$L_{dif} = \left( (H'_{con})^T H'_c \right)^2 + \left( (H'_{sy})^T H'_c \right)^2 \quad (19)$$

Combining the above regularization and loss functions, the overall objective function is obtained, and the parameters of the model are optimized by backpropagation as follows:

$$L = - \sum_{i \in S} \sum_{j \in P} y_i^j \log p_i^j + \lambda_1 L_{sim} + \lambda_2 L_{dif} + \lambda_3 (R_O + R_D) + \lambda_4 \|\theta\|_2 \quad (20)$$

where  $S$  is the sample of all input sentences;  $P$  is all possible sentiment polarities;  $y$  is the true label;  $p$  is the value predicted by the model;  $\lambda_1$  and  $\lambda_2$  are the coefficients of the two loss function terms;  $\lambda_3$  and  $\lambda_4$  are the regularization coefficients;  $\theta$  is the trainable parameter.

## 4 Experiments

In this section, we first introduce the public standard datasets for verification and the experimental details for implementation. Secondly, we introduce the baseline models for comparison, then we describe the experimental results and analysis, as well as the ablation experiments and case studies. Finally, we discuss the effects of the number of GCN layers and different types of labels on the experimental results.

We use the PyTorch development framework to implement MultiGCN, which is deployed on two NVIDIA GeForce RTX3090 GPU, the version of CUDA is 11.0.

### 4.1 Datasets

We evaluate our proposed model on four benchmark datasets, including the Restaurant and Laptop datasets from the SemEval 2014 ABSA challenge [44], the Twitter dataset from Dong et al. [45], and the MAMS dataset released recently by Jiang et al. [46]. Each sentence in all four datasets is marked by three sentiment polarities (positive, negative, and neutral) according to the given aspect. The statistics of the four datasets are indicated in Table 1.

### 4.2 Implementation detail

In this work, we use the LAL parser [47] to obtain the dependency tree. To initialize the word embeddings in our experiments, we exploit pretrained 300-dimensional Glove vectors [48]. Moreover, before the BiLSTM encoder, the relative position embeddings of the aspect words as well as the part-of-speech embeddings of each word in the sentences are concatenated with the word embeddings, and both dimensions are set to 30. The hidden layer dimension is set to 50 during propagation. Our model is optimized using the Adam optimizer with a learning rate of 0.001. To prevent overfitting, the L2 regularization is set to 0.0001, and the dropout rate on input word embeddings and GCN modules are 0.7 and 0.1, respectively. After that, we set up 50 epochs for training, with a batch size of 16.

We use two evaluation metrics: accuracy and Macro Average F1. The former is the most common performance indicator for classification, and the latter is more suitable for datasets with imbalanced classes.

**Table 1** Statistics of datasets

Dataset	Division	Positive	Negative	Neutral
Restaurant	Training	2,164	807	637
	Testing	727	196	196
Laptop	Training	976	851	455
	Testing	337	128	167
Twitter	Training	1,507	1,528	3,016
	Testing	172	169	336
MAMS	Training	3,380	2,764	5,042
	Testing	400	329	607

### 4.3 Baseline models

Our proposed model will be compared with several state-of-the-art baseline models to demonstrate the effectiveness of MultiGCN. A brief description of the baseline models is shown below.

- 1) ATAE-LSTM [8] combines attention with LSTM to solve the aspect level sentiment analysis problem by using attention to obtain contextual information that is more important to different aspects.
- 2) IAN [9] first models target and context separately, and then links them using an attention mechanism to obtain the representation that incorporates information about their interaction.
- 3) AOA [10] learns aspect representations and sentence representations that explicitly capture the interaction between the aspect and context, and automatically focuses on the important parts of the sentence.
- 4) MGAN [11] can capture word-level interactions between the aspect and context, alleviating information loss in a coarse-grained attention mechanism.
- 5) ASGCN [22] uses the syntax dependency structure in sentences and solves the problem of long-distance word dependency in ABSA.
- 6) CDT [23] enhances embeddings by using GCNs that act directly on the sentence dependency tree to better learn sentence representations.
- 7) kumaGCN [24] uses automatically induced aspect-specific graphs to obtain more comprehensive syntactic features.
- 8) BiGCN [42] fully integrates hierarchical syntactic graphs and lexical graphs and obtains aspect-oriented representations through the mask and gating mechanism.
- 9) InterGCN [25] can simultaneously obtain important aspect-focused and inter-aspect information by using heterogeneous graphs.
- 10) CL-GCN [49] makes use of both global and local structural information and uses attention mechanisms to fuse the two kinds of information.

- 11) RGAT [27] uses dependent label information and a new attention function to help targets better capture useful contextual information.

### 4.4 Experimental results and analysis

Table 2 presents the experimental results for all baseline models and our proposed MultiGCN model using accuracy and Macro-F1 as evaluation metrics.

Our model achieves optimal results on the Restaurant, Laptop, Twitter and MAMS datasets. Among them, MAMS has the largest amount of data, in the meanwhile, MultiGCN has the largest improvement over other state-of-the-art models on the MAMS dataset, with both accuracy and Macro-F1 increasing by 1.86 percent, which demonstrates the potential of our model on large datasets. Similarly, on the Restaurant, Laptop and Twitter datasets, accuracy improved by 0.27 percent, 0.78 percent and 0.27 percent, respectively, and Macro-F1 improved by 1.11 percent, 0.97 percent and 0.15 percent, respectively. This demonstrates the universality of our approach and the ability to take advantage of structural, contextual, and common information in sentences.

Furthermore, comparing all models, it can be observed that models using a dependency tree or its variants achieve higher accuracy rates, indicating that the use of structural information is beneficial for the ABSA task. However, considering only structural information has become increasingly limited. On the one hand, model optimization can be achieved with the expansion of structural information, e.g., by increasing the types of dependencies between words. On the other hand, other information about the sentence, such as content information between contexts, can be considered interactively.

### 4.5 Ablation study

As shown in Table 3, ablation experiments were conducted on the Restaurant, Laptop, Twitter, and MAMS datasets to demonstrate the validity of the components of our model. Four main components are ablated: dependency label,

**Table 2** Experimental results comparison on four publicly available datasets

Models	Restaurant		Laptop		Twitter		MAMS	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
ATAE-LSTM	78.60	67.02	68.88	63.93	70.14	66.03	–	–
IAN	79.26	70.09	72.05	67.38	72.50	70.81	–	–
AOA	79.97	70.42	72.62	67.52	72.30	70.20	–	–
MGAN	81.25	71.94	75.39	72.47	72.54	70.81	–	–
ASGCN	80.77	72.02	75.55	71.05	72.15	70.40	–	–
CDT	82.30	74.02	77.19	72.99	74.66	73.66	80.70	79.79
kumaGCN	81.43	73.64	76.12	72.42	72.45	70.77	–	–
BiGCN	81.97	73.48	74.59	71.84	74.16	73.35	–	–
InterGCN	82.23	74.01	77.86	74.32	–	–	–	–
CL-GCN	82.11	73.46	76.91	72.76	73.26	71.26	–	–
RGAT	83.55	75.99	78.02	74.00	75.36	74.15	81.75	80.87
Our MultiGCN	<b>83.82</b>	<b>77.10</b>	<b>78.80</b>	<b>74.97</b>	<b>75.63</b>	<b>74.30</b>	<b>83.61</b>	<b>82.73</b>

The best results on each dataset are highlighted in bold font

common information extraction module, fusion mechanism, and improved loss function.

MultiGCN w/o DL indicates that our model removes the dependency label. The results show that this is the most significant factor causing a decrease in performance on the Restaurant, Laptop and MAMS datasets, and a degree of decrease in both accuracy and Macro-F1 on the Twitter dataset. It is suggested that the introduction of dependency labels can clarify the relationship between aspect words and context, thus locating the opinion words with sentiment polarity more precisely and improving the accuracy of prediction results.

MultiGCN w/o ComInfo means that the model removes the common information extraction module. Compared with the complete MultiGCN, performance on all datasets has declined, while has declined most dramatically in the Twitter dataset. We believe it is due to the fact that the Twitter dataset has both well-structured sentences and many colloquial expressions. Therefore, the dataset is more

sensitive to the information that combines structure and content.

MultiGCN w/o FM means that we remove the fusion mechanism from the model. In other words, instead of exchanging structural information, contextual information, and common information containing both, the final generated sentence expression is fed directly into the pooling layer. Owing to the lack of interaction between features, it can be seen that the performance decreases to varying degrees on all datasets.

Finally, experiments are conducted for our modified loss function. MultiGCN w/o  $L_{dif}$  refers to the model that removes the difference loss during training, i.e., the difference between features containing both types of information and the structure or content information alone is not considered. MultiGCN w/o  $L_{sim}$  indicates that the model is trained with similarity loss removed, i.e., without considering the correlation between structure information and content information. The experimental results show that

**Table 3** Ablation study on the four datasets

Models	Restaurant		Laptop		Twitter		MAMS	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
MultiGCN	83.82	77.10	78.80	74.97	75.63	74.30	83.61	82.73
MultiGCN w/o DL	82.66	75.81	77.06	72.59	74.74	73.5	82.26	81.59
MultiGCN w/o ComInfo	83.38	76.29	78.16	74.75	74.15	72.82	82.63	82.04
MultiGCN w/o FM	82.81	75.88	77.18	73.22	75.18	73.31	83.11	82.31
MultiGCN w/o $L_{dif}$	82.93	75.80	77.85	73.96	74.74	73.57	83.38	82.28
MultiGCN w/o $L_{sim}$	83.29	75.45	77.37	73.44	75.04	74.16	83.08	82.11
MultiGCN w/o $L_{dif}$ & $L_{sim}$	83.11	74.79	78.01	74.62	75.04	73.33	83.23	82.29



MultiGCN w/o ComInfo	0.6576	0.8582	0.6027	0.8272	0.9489	0.4967	0.4228	0.0086	0.7533	0.8984	1	0.2505
MultiGCN w/o DL	0.7268	0.6835	0.681	0.1233	0.0123	0.0285	0.7548	0.733	0.0311	0.2552	1	0.1552
MultiGCN	0.1165	0.1391	0.0383	0.0414	0.0944	0.0963	0.0363	0.0771	0.551	1	0.5431	0.0594
	The	pizza	is	the	best	if	you	like	thin	crusted	pizza	.

Fig. 4 Visualization of word relevance scores

similarity loss has a greater impact on most datasets. We reason that this is because the vast majority of the data have both syntax structure and context information. Thus, if the two features can be well combined, the model training can be better completed. MultiGCN w/o  $L_{dif}$  &  $L_{sim}$  denotes that both the difference loss and the similarity loss are removed. As can be seen, combining the two loss functions still facilitates the overall training of the model, even with better results.

### 4.6 Case study

To further analyze the MultiGCN model, we investigate the effect of different components of the model visually through a case study. We adopt the mask method proposed by Sun et al. [23] to calculate the contribution of different words in a sentence to the final emotional prediction. The larger the value, the more influential the final judgment. The calculation is given as follows:

$$\gamma(w, s) = \frac{1}{m} |h_s - h_{s/w}| \tag{21}$$

$$m = \max_{w \in s} \gamma(w, s) \tag{22}$$

where  $s$  denotes the example sentence we selected, i.e., “The pizza is the best if you like thin crusted pizza.”;  $w$  denotes the word in the sentence;  $h_s$  denotes the final representation of the example sentence output by the model;  $h_{s/w}$  denotes a sentence representation that masks out one word from front to back.

In the example sentence, the aspect word is “thin crusted pizza”, and sentiment polarity is neutral. Figure 4 visualizes the relevance scores of the words in the sentence as a heat map, which shows that our model MultiGCN can accurately determine the contextual information that is most relevant to the aspect words and minimize the focus on irrelevant information. In addition, with the removal of the dependency label or common information extraction module, the scores for each word were changed, and the focus on irrelevant context or even distracting information was increased. In contrast, the common information extraction module had a greater impact on the whole sentence. After removing this module, the score for interfering information in the whole sentence representation increases significantly and is almost identical to the relevant information, which can easily cause errors in judgment. Thus, it confirms that our proposed model can accurately locate the opinion words and correctly predict the sentiment polarity of sentences.

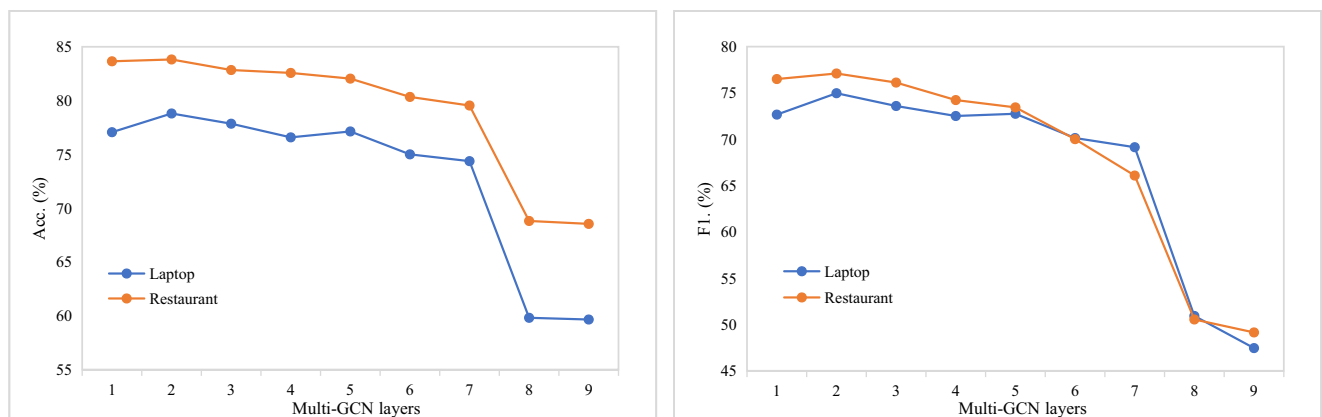
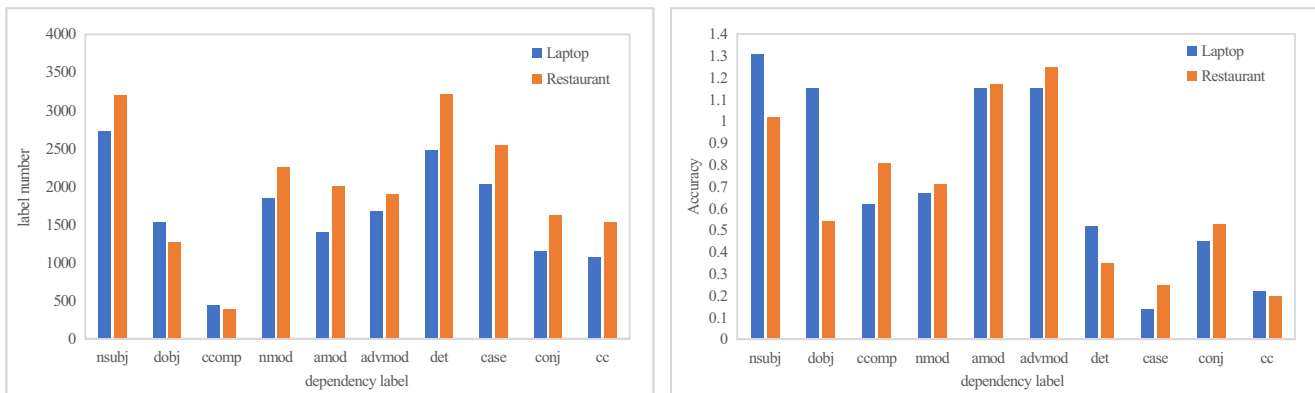


Fig. 5 Effect of GCN layers on accuracy and Macro-F1



**Fig. 6** Number of dependency labels and the impact on the accuracy

#### 4.7 Performance of the multi-GCN layer number

To analyze the effect of the number of GCN layers on the experimental results, we select the Laptop and Restaurant datasets and conduct experiments on them for GCN layers from 1 to 9. The results are shown in the form of a line graph, as shown in Fig. 5. It is proved that the model works best when the number of GCN layers is 2, i.e., both accuracy and Macro-F1 are the highest. However, if the number of GCN layers increases to 8, the accuracy and Macro-F1 greatly decreases, which is attributed to overfitting caused by the high complexity of the model. In summary, our model adopts a 2-layer GCN on all datasets and achieves good results.

#### 4.8 Effects of different dependency labels

To further explore the effect of dependency labels on accuracy, we choose the 10 most common labels and remove them one by one to study the effect of the decrease in accuracy for different labels on the Laptop and Restaurant datasets. The results are illustrated in Fig. 6.

Here are some observations. First, on the Laptop dataset, nsubj, dobj, amod and advmod labels have a great impact on the accuracy. We believe this is because these tags belong to nouns and noun modifications, which are crucial for

measuring the sentiment of sentences. Therefore, deleting these tags also has the greatest impact. In addition, ccomp, nmod, and det labels play a certain role, resulting in an average decrease in accuracy of 0.7. The rest of the labels have little effect on the model, mainly those indicating prepositions or conjunctions, such as conj and cc labels. Second, on the Restaurant dataset, most of the results are consistent with the Laptop dataset. However, as for the dobj tag, the impact on the results is much smaller than on the Laptop dataset. We speculate that it is because the dobj label only accounts for 35 percent of the Restaurant dataset, while it accounts for 67 percent in the Laptop dataset, thus it has a relatively low impact on model accuracy.

#### 4.9 Analysis of model efficiency

To further evaluate the training efficiency of the model, we compare MultiGCN with RGAT in terms of the average training time per epoch and the total number of parameters, as shown in Table 4. We unify the batch size and hidden size of the two models to ensure a fair comparison. The results show that the average training time for epoch is slightly longer than that of the baseline model, due to the additional operations such as the fusion mechanism and the optimization of the loss function, but the overall time difference with the baseline model is not significant.

**Table 4** Efficiency comparison between MultiGCN and RGAT

Models	Restaurant		Laptop		Twitter		MAMS	
	Times (s)	Params	Times (s)	Params	Times (s)	Params	Times (s)	Params
MultiGCN	42.65	1.58M	19.64	1.28M	50.38	4.16M	106.34	2.74M
RGAT	34.01	1.91M	19.55	1.49M	48.21	4.50M	95.07	2.96M

For the total number of parameters, MultiGCN has fewer parameters since our method used GCN to encode the features.

## 5 Conclusion and future work

In this paper, we propose a MultiGCN model for the ABSA task. The different types of GCN, the common information extraction module, and the fusion mechanism are used to properly integrate the structure of the sentence with the content information. With the improved loss function, the parameters of the model are continuously optimized. The experimental results on the four benchmark datasets show that our model achieves state-of-the-art performance. In future work, we will consider applying external knowledge to help the proposed model understand the relationship between words in a sentence, thereby improving its performance.

**Acknowledgements** This research is supported by the National Natural Science Foundation of China (62077027), the Ministry of Science and Technology of the People's Republic of China (2018YFC2002500), the Jilin Province Development and Reform Commission, China (2019C053-1), the Education Department of Jilin Province, China (JJKH20200993K), the Department of Science and Technology of Jilin Province, China (20200801002GH), and the European Union's Horizon 2020 FET Proactive project "WeNet-The Internet of us" (No. 823783).

## Declarations

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

## References

- Zhang M, Zhang Y, Vo D-T (2016) Gated neural networks for targeted sentiment analysis. In: Thirtieth AAAI conference on artificial intelligence
- Wang S, Mazumder S, Liu B, Zhou M, Chang Y (2018) Target-sensitive memory networks for aspect sentiment classification. In: Proceedings of the 56th annual meeting of the association for computational linguistics (vol 1: long papers), pp 957–967
- Nazir A, Rao Y, Wu L, Sun L (2020) Issues and challenges of aspect-based sentiment analysis: a comprehensive survey, *IEEE Trans Affect Comput*, pp 1–1
- Liang B, Su H, Gui L, Cambria E, Xu R (2022) Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowl Based Syst* 107643
- Ren F, Feng L, Xiao D, Cai M, Cheng S (2020) Dnet: a lightweight and efficient model for aspect based sentiment analysis. *Expert Syst Appl*, pp 113393
- Tang D, Qin B, Feng X, Liu T (2016) Effective lstm for target-dependent sentiment classification. In: Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers, pp 3298–3307
- Tang D, Qin B, Feng X, Liu T (2016) Effective LSTMs for target-dependent sentiment classification. In: Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers, pp 3298–3307
- Wang Y, Huang M, Zhu X, Zhao L (2016) Attention-based lstm for aspect-level sentiment classification. In: Proceedings of the 2016 conference on empirical methods in natural language processing, pp 606–615
- Ma D, Li S, Zhang X, Wang H (2017) Interactive attention networks for aspect-level sentiment classification. In: Proceedings of the twenty-sixth international joint conference on artificial intelligence, IJCAI-17, pp 4068–4074 . <https://doi.org/10.24963/ijcai.2017/568>
- Huang B, Ou Y, Carley KM (2018) Aspect level sentiment classification with attention-over-attention neural networks. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation, pp 197–206
- Fan F, Feng Y, Zhao D (2018) Multi-grained attention network for aspect-level sentiment classification. In: Proceedings of the 2018 conference on empirical methods in natural language processing, pp 3433–3442 . <https://doi.org/10.18653/v1/D18-1380>
- Yadav RK, Jiao L, Goodwin M, Grammo O-C (2021) Positionless aspect based sentiment analysis using attention mechanism. *Knowl Based Syst*, pp 107136
- Liu M, Zhou F, Chen K, Zhao Y (2021) Co-attention networks based on aspect and context for aspect-level sentiment analysis. *Knowl Based Syst*, pp 106810
- Huang B, Carley K (2018) Parameterized convolutional neural networks for aspect level sentiment classification. In: Proceedings of the 2018 conference on empirical methods in natural language processing, pp 1091–1096. <https://doi.org/10.18653/v1/D18-1136>
- Fan C, Gao Q, Du J, Gui L, Xu R, Wong K-F (2018) Convolution-based memory network for aspect-based sentiment analysis. In: The 41st international ACM SIGIR conference on research & development in information retrieval, pp 1161–1164
- Li X, Bing L, Lam W, Shi B (2018) Transformation networks for target-oriented sentiment classification. In: Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: long papers), pp 946–956
- Zhang Y, Qi P, Manning CD (2018) Graph convolution over pruned dependency trees improves relation extraction. In: Proceedings of the 2018 conference on empirical methods in natural language processing, pp 2205–2215
- Huang L, Sun X, Li S, Zhang L, Wang H (2020) Syntax-aware graph attention network for aspect-level sentiment classification. In: Proceedings of the 28th international conference on computational linguistics, pp 799–810
- Zhang C, Li Q, Song D (2019) Syntax-aware aspect-level sentiment classification with proximity-weighted convolution network. In: Proceedings of the 42nd international ACM SIGIR conference on research and development in information retrieval, pp 1145–1148
- Phan MH, Ogunbona PO (2020) Modelling context and syntactical features for aspect-based sentiment analysis. In: Proceedings of the 58th annual meeting of the association for computational linguistics, pp 3211–3220
- Tian Y, Chen G, Song Y (2021) Enhancing aspect-level sentiment analysis with word dependencies. In: Proceedings of the 16th conference of the european chapter of the association for computational linguistics: Main volume, pp 3726–3739
- Zhang C, Li Q, Song D (2019) Aspect-based sentiment classification with aspect-specific graph convolutional networks. In: EMNLP-IJCNLP, pp 4568–4578. <https://doi.org/10.18653/v1/D19-1464>

23. Sun K, Zhang R, Mensah S, Mao Y, Liu X (2019) Aspect-level sentiment analysis via convolution over dependency tree. In: Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), pp 5679–5688
24. Chen C, Teng Z, Zhang Y (2020) Inducing target-specific latent structures for aspect sentiment classification. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), pp 5596–5607
25. Liang B, Yin R, Gui L, Du J, Xu R (2020) Jointly learning aspect-focused and inter-aspect relations with graph convolutional networks for aspect sentiment analysis. In: Proceedings of the 28th international conference on computational linguistics, pp 150–161
26. Cai H, Tu Y, Zhou X, Yu J, Xia R (2020) Aspect-category based sentiment analysis with hierarchical graph convolutional network. In: Proceedings of the 28th international conference on computational linguistics, pp 833–843
27. Bai X, Liu P, Zhang Y (2021) Investigating typed syntactic dependencies for targeted sentiment classification using graph attention neural network. *IEEE/ACM Trans Audio Speech Lang Process*, pp 503–514. <https://doi.org/10.1109/TASLP.2020.3042009>
28. Huang B, Carley K (2019) Syntax-aware aspect level sentiment classification with graph attention networks. In: Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), pp 5469–5477. <https://doi.org/10.18653/v1/D19-1549>
29. Wang K, Shen W, Yang Y, Quan X, Wang R (2020) Relational graph attention network for aspect-based sentiment analysis. In: Proceedings of the 58th annual meeting of the association for computational linguistics, pp 3229–3238
30. Yuan L, Wang J, Yu L-C, Zhang X (2020) Graph attention network with memory fusion for aspect-level sentiment analysis. In: Proceedings of the 1st conference of the asia-pacific chapter of the association for computational linguistics and the 10th international joint conference on natural language processing, pp 27–36
31. Liang B, Su H, Gui L, Cambria E, Xu R (2022) Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowl Based Syst*, pp 107643
32. Zhao P, Hou L, Wu O (2020) Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification. *Knowl Based Syst*, pp 105443
33. Sun C, Huang L, Qiu X (2019) Utilizing BERT for aspect-based sentiment analysis via constructing auxiliary sentence. In: Proceedings of the 2019 conference of the north american chapter of the association for computational linguistics: Human language technologies, vol 1. (long and short papers), pp 380–385. <https://doi.org/10.18653/v1/N19-1035>
34. Hou X, Huang J, Wang G, Qi P, He X, Zhou B (2021) Selective attention based graph convolutional networks for aspect-level sentiment classification. In: Proceedings of the Fifteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-15), pp 83–93
35. Hou X, Qi P, Wang G, Ying R, Huang J, He X, Zhou B (2021) Graph ensemble learning over multiple dependency trees for aspect-level sentiment classification. In: Proceedings of the 2021 conference of the north american chapter of the association for computational linguistics: Human language technologies, pp 2884–2894
36. Tang J, Lu Z, Su J, Ge Y, Song L, Sun L, Luo J (2019) Progressive self-supervised attention learning for aspect-level sentiment analysis. In: Proceedings of the 57th annual meeting of the association for computational linguistics, pp 557–566
37. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 25:1097–1105
38. Sukhbaatar S, Weston J, Fergus R et al (2015) End-to-end memory networks. *Adv Neural Inf Process Syst*, vol 28
39. Tang D, Qin B, Liu T (2016) Aspect level sentiment classification with deep memory network. In: Proceedings of the 2016 conference on empirical methods in natural language processing, pp 214–224
40. Chen P, Sun Z, Bing L, Yang W (2017) Recurrent attention network on memory for aspect sentiment analysis. In: Proceedings of the 2017 conference on empirical methods in natural language processing, pp 452–461
41. Li C, Guo X, Mei Q (2017) Deep memory networks for attitude identification. In: Proceedings of the tenth ACM international conference on web search and data mining, pp 671–680
42. Zhang M, Qian T (2020) Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis. In: Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP), pp 3540–3549
43. Li R, Chen H, Feng F, Ma Z, Wang X, Hovy E (2021) Dual graph convolutional networks for aspect-based sentiment analysis. In: Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (Volume 1: long papers), pp 6319–6329
44. Pontiki M, Galanis D, Pavlopoulos J, Papageorgiou H, Manandhar S (2014) Semeval-2014 task 4: Aspect based sentiment analysis proceedings of international workshop on semantic evaluation at
45. Dong L, Wei F, Tan C, Tang D, Zhou M, Xu K (2014) Adaptive recursive neural network for target-dependent twitter sentiment classification. In: Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short Papers), pp 49–54
46. Jiang Q, Chen L, Xu R, Ao X, Yang M (2019) A challenge dataset and effective models for aspect-based sentiment analysis. In: Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP), pp 6280–6285
47. Mrini K, Démoncourt F, Tran QH, Bui T, Chang W, Nakashole N (2020) Rethinking self-attention: Towards interpretability in neural parsing. In: Findings of the association for computational linguistics: EMNLP 2020, pp 731–742. <https://doi.org/10.18653/v1/2020.findings-emnlp.65>
48. Pennington J, Socher R, Manning CD (2014) Glove: Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543
49. Zhu X, Zhu L, Guo J, Liang S, Dietze S (2021) GI-gcn: Global and local dependency guided graph convolutional networks for aspect-based sentiment classification. *Expert Syst Appl*

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.



**Yuting Ma** is currently studying for a M.S. degree in the College of Software, Jilin University. She received the B.S. degree in the College of Software, Taiyuan University of Technology. Her main research interests include natural language processing and Sentiment Analysis.



**Qiang Shen** is a Ph.D student at the College of computer science and technology in Jilin University, China and a joint-training Ph.D student at Department of Information Engineering and Computer Science of the University of Trento, Italy. His research interests include federated learning, transfer learning, human activity recognition, and time series classification.



**Rui Song** is currently studying for a Ph.D. degree in the School of Artificial Intelligence, Jilin University. He received the B.S. degree from the College of Software of Jilin University. His main research interests include natural language processing and graph neural networks.



**Hao Xu** received a Ph.D. degree in University of Trento, Italy. He is currently a full professor, College of Computer Science and Technology, Jilin University. His main research interests include knowledge graph and intelligent interaction.



**Xue Gu** is currently studying for her PhD in the School of Electrical and Computer Engineering, Minho University. She obtained her B.S. from the College of Computer Science and Technology, Jilin University. Her main research interests include Natural Language Processing and Sentiment Analysis.

## Affiliations

Yuting Ma<sup>1</sup> · Rui Song<sup>2</sup> · Xue Gu<sup>3</sup> · Qiang Shen<sup>4</sup> · Hao Xu<sup>4</sup>

Yuting Ma  
mayt20@mails.jlu.edu.cn

Rui Song  
songrui20@mails.jlu.edu.cn

Xue Gu  
id9267@alunos.uminho.pt

Qiang Shen  
shenqiang19@mails.jlu.edu.cn

- <sup>1</sup> College of Software, Jilin University, Changchun, 130012, China
- <sup>2</sup> School of Artificial Intelligence, Jilin University, Changchun, 130012, China
- <sup>3</sup> Department of Industrial Electronics, University of Minho, Guimarães, 4800-058, Portugal
- <sup>4</sup> College of Computer Science and Technology, Jilin University, Changchun, 130012, China