



OPEN ACCESS

EDITED BY

Xin Lu,
De Montfort University, United Kingdom

REVIEWED BY

E. Zhang,
University of Leicester, United Kingdom
Yinong Chen,
Arizona State University, United States
Amin Ul Haq,
University of Electronic Science and
Technology of China, China

*CORRESPONDENCE

Qianhua Cai,
✉ caiqianhua@m.scnu.edu.cn

RECEIVED 11 November 2023

ACCEPTED 11 March 2024

PUBLISHED 05 April 2024

CITATION

Huang Y, Dai A, Cao S, Kuang Q, Zhao H and
Cai Q (2024), A dual contrastive learning-based
graph convolutional network with syntax label
enhancement for aspect-based
sentiment classification.
Front. Phys. 12:1336795.
doi: 10.3389/fphy.2024.1336795

COPYRIGHT

© 2024 Huang, Dai, Cao, Kuang, Zhao and Cai.
This is an open-access article distributed under
the terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

A dual contrastive learning-based graph convolutional network with syntax label enhancement for aspect-based sentiment classification

Yuyan Huang^{1,2}, Anan Dai³, Sha Cao⁴, Qihua Kuang^{1,5},
Hongya Zhao⁶ and Qianhua Cai^{1*}

¹Department of Electronics and Information Engineering, South China Normal University, Foshan, China, ²Datastory, Guangzhou, China, ³China Merchants Bank Foshan Branch, Foshan, China, ⁴Department of Financial Mathematics, University of Chicago, Chicago, IL, United States, ⁵Guangzhou Qizhi Information Technology Co., Ltd., Guangzhou, China, ⁶Industrial Centre School of Undergraduate Education, Shenzhen Polytechnic University, Shenzhen, China

Introduction: Aspect-based sentiment classification is a fine-grained sentiment classification task. State-of-the-art approaches in this field leverage graph neural networks to integrate sentence syntax dependency. However, current methods fail to exploit the data augmentation in encoding and ignore the syntactic relation in sentiment delivery.

Methods: In this work, we propose a novel graph neural network-based architecture with dual contrastive learning and syntax label enhancement. Specifically, a contrastive learning-based contextual encoder is designed, integrating sentiment information for semantics learning. Moreover, a weighted label-enhanced syntactic graph neural network is established to use both the syntactic relation and syntax dependency, which optimizes the syntactic weight between words. A syntactic triplet between words is generated. A syntax label-based contrastive learning scheme is developed to map the triplets into a unified feature space for syntactic information learning.

Results: Experiments on five publicly available datasets show that our model substantially outperforms the baseline methods.

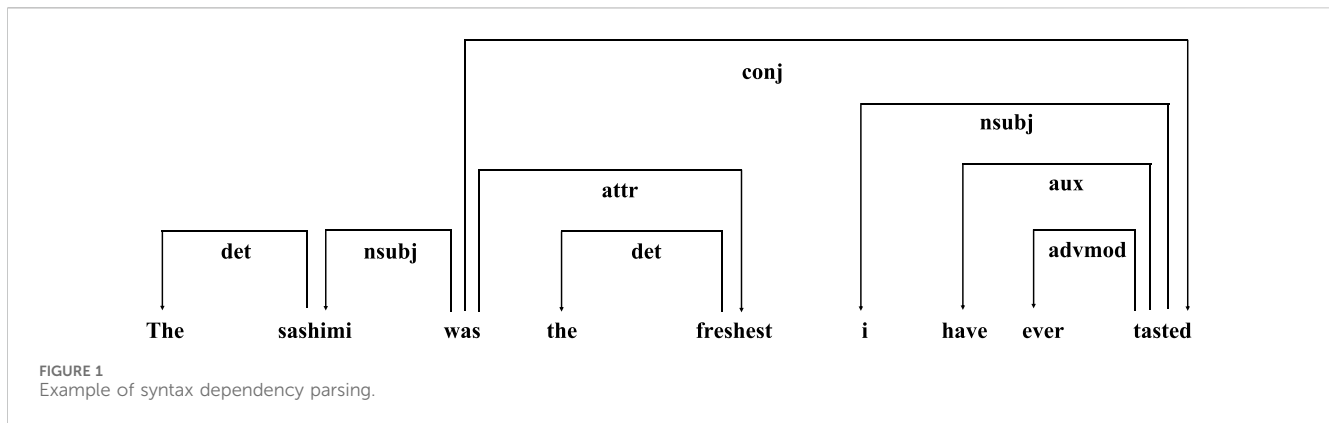
Discussion: As such, the proposed method shows its effectiveness in aspect-based sentiment classification tasks.

KEYWORDS

aspect-based sentiment classification, graph convolutional networks, dual contrastive learning, syntax label enhancement, bidirectional encoder representations from transformers (BERT)

1 Introduction

Aspect-based sentiment classification (ABSC) is a fundamental task in sentiment analysis [1]; [2], which aims to infer the sentiment of a specific aspect in sentences [3]. Generally, the sentiment of each aspect is classified according to a predefined set of sentiment polarities, i.e., positive, neutral, or negative. For example, in the comment “the price is reasonable, although the service is poor,” the sentiment toward aspects “price” and “service” is positive and negative, respectively.



In general, an ABSC process involves two steps: the identification of sentiment information toward the aspect from the context and the classification of the expressed sentiment from predefined sentiment polarities [4]. Comprehensively, the first step contains key contextual information learning and aspect–context word relation establishment. To capture important contextual words and prevent redundant information, recent publications reveal that encoders and attention networks are taken to encode the sequential information and determine the attentive weights of contexts, respectively [5]. Typically, these deep learning methods are trained via a large amount of textual data to improve their working performance. Notwithstanding, the existing manually annotated data resources are still limited, which causes issues such as model overfitting. As a result, the precise capturing of key contextual words remains challenging. More recently, contrastive learning shows its superiority under the condition of limited training samples. Based on data augmentation, both positive and negative samples are generated. By setting contrastive loss of training models, the representations of positive samples are brought closer, while those of negative samples are pushed apart. In line with the contrastive learning, the model training can be improved, which paves a way for key contextual information learning in ABSC tasks.

On the basis of key contextual information, the aspect–opinion word relation mainly lies in syntax dependency of the sentence [6]. With the parsing of syntax dependency, the relation between the aspect and context words is built. Ongoing studies substantially focus solely on the distance of words while neglecting the syntax label of specific context words toward the aspect. That is, all syntactic relations are interpreted as the same. Figure 1 shows the syntax structure of a given sentence. The establishment of the subject–predicate syntactic relation (nsubj) and adjective modifier syntactic relation (attr) plays a dominate role in sentiment classification, especially compared with other syntactic information. Moreover, the syntax label is also the foundation of textual logical reasoning due to its effects in distinguishing the importance among syntactic relations. So much is the significance of the syntax label that it can be further applied to the aspect–opinion word relation establishment in ABSC.

To address the above issues in ABSC, we propose a graph convolutional network (GCN) based on dual contrastive learning and syntax label enhancement (i.e., DCL-GCN). First, a contrastive learning-based encoder is devised, which brings the context representations of the same sentiment closer and pushes those of different sentiments apart. Furthermore, a weighted label-enhanced syntactic GCN is put forward, dealing with not only the syntactic

relation but also the syntax dependencies among words. Lastly, a contrastive learning scheme that focuses on the sentence syntax label is developed. A syntactic triplet between words is constructed. The same syntax label-based triplets are given similar semantic representations, while different syntax label-based triplets are distinguished. Thereby, the syntax and semantics are integrated, which contributes to the sentiment classification.

The contribution of our work is three-fold and given as follows:

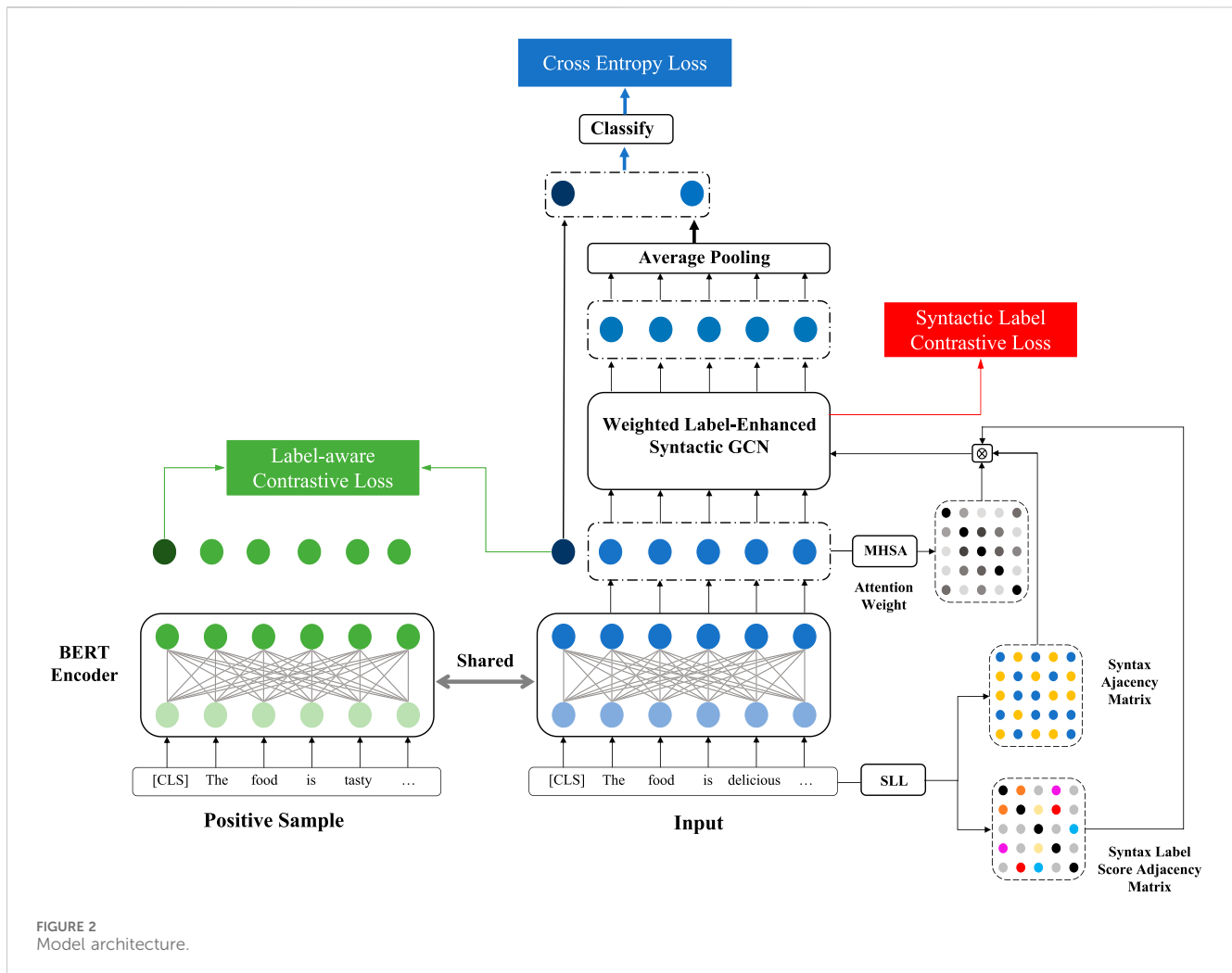
- A GCN-based ABSC method is proposed with the integration of dual contrastive learning and syntax label enhancement. Specifically, the sentence is encoded using contrastive learning to bring the context representations of the same sentiment closer and push those of different sentiments apart.
- A weighted label-enhanced syntactic GCN and a contrastive learning scheme are established to tackle the sentence syntax. A syntactic triplet between words can be generated. The same syntax label-based triplets are given similar semantic representations to facilitate the ABSC.
- Experiments conducted on five benchmark datasets demonstrate that our model achieves state-of-the-art results. The proposed method significantly improves the working performance compared to competitive baselines in the ABSC field.

2 Related work

Owing to the advancement of deep learning networks, current methods with various structures are widely developed, aiming to identify their superiority in ABSC tasks [7]. ABSC models are devised to deal with either semantics [8], syntax [9], or both [10] from the given text. In this section, these two major issues in the field of ABSC are presented. In order to achieve better working performance, previous work and their findings about these two focuses are dedicatedly investigated and depicted.

2.1 Contextual information learning

One bottleneck in ABSC comes from capturing key contextual words, which considerably affects the aspect–opinion word relation modeling. Much recent work uses neural networks, attention networks, or both to concentrate on useful contextual information [11]; [12]. Tang



et al. focused on different contextual parts based on LSTM, targeting at obtaining valuable information [13]. In addition, attention-based neural networks are proposed to discriminate more relevant features toward the aspect [14]; [15]. Sun et al. used a BERT-based model to capture semantic features from contexts via fine-tuning, which significantly improves the working performance [16]. Text encoders are widely applied to various tasks [17,18]. Encouragingly, advances in contrastive learning hold great potential in natural language processing (NLP) tasks. Suresh et al. integrated contrastive learning strategy into the pre-training of Bidirectional Encoder Representations from Transformers (BERT) to improve the model efficacy [19]. A contrastive loss among different input categories is introduced, while a weight network refines the differences between each sample pair. In our work, contrastive learning can be taken to distinguish the contextual representations during sentence encoding.

2.2 Syntax dependency parsing

The parsing of syntax dependency plays a pivotal role in the field of ABSC due to its relation establishment between the aspect and contextual words. Previous work primarily tackles the syntactic relation of either single or multiple word pairs. In recent years,

the application of a GCN in NLP gave rise to new opportunities in a number of fields [20]; [21]. Regarding sentence syntax parsing, Sun et al. transformed the syntax dependency into an adjacency matrix and propagated the syntactic information using the GCN [22]. Furthermore, Zhang et al. incorporated the aspect-oriented attention mechanism to benefit the contextual information extraction toward a specific aspect [23]. To extract both aspect-focused and inter-aspect sentiment information, an interactive graph convolutional network (InterGCN) is built to leverage the sentiment dependencies of the context [24]. Wang et al. reconstructed the syntax dependency tree rooted at an aspect. A relational graph attention network (R-GAT) is then proposed to encode the aspect-oriented dependency tree and to establish the syntactic relation between the aspect and its opinion words [25].

3 Methodology

A dual contrastive learning GCN (DCL-GCN) is devised on the task of ABSC. Figure 2 shows the framework of the proposed model. A pretrained BERT model is used as the sentence encoder. A contrastive learning scheme is incorporated into contextual encoding during model training, which enhances the semantic information via sentiment labels

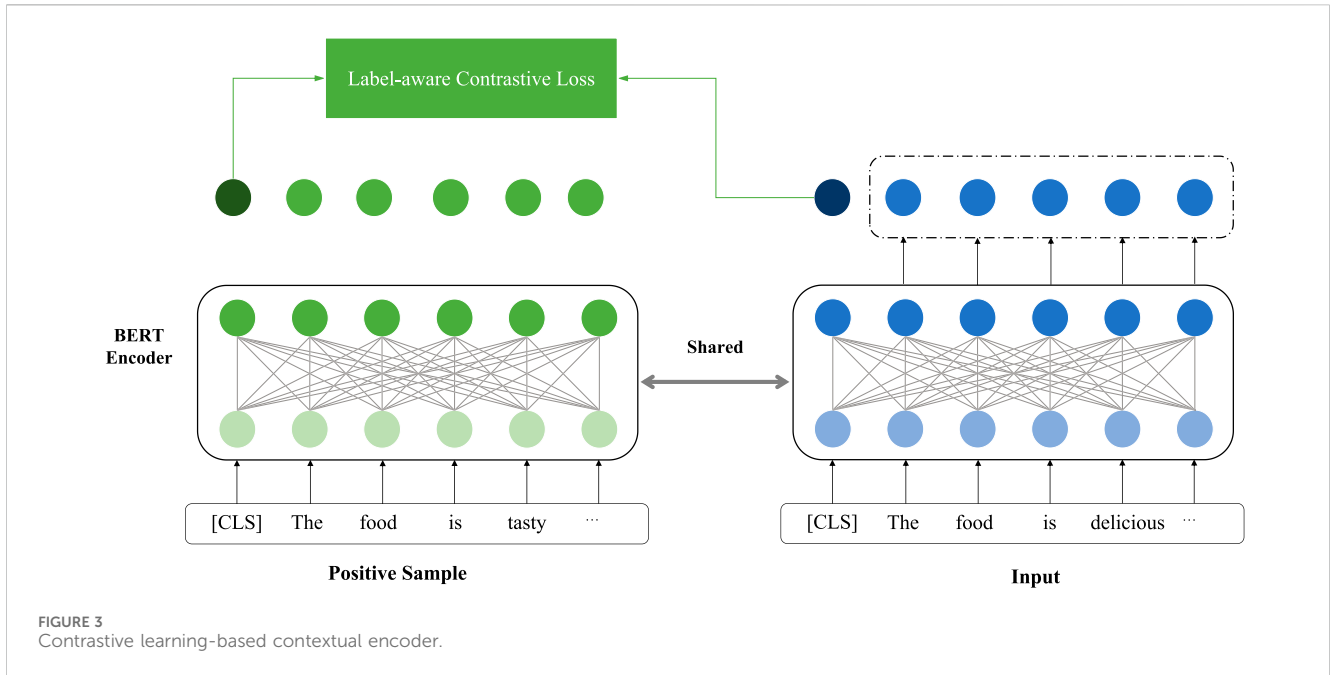


FIGURE 3 Contrastive learning-based contextual encoder.

to obtain differentiated contextual representations. Then, both the semantic and syntactic features are integrated within a weighted label GCN, aiming at addressing the syntactic relation of context words with the aspect. In line with contrastive learning, the syntax labels of words are used for learning the sentence syntax at a higher level. The sentiment polarity is predicted by sending the final sentence representation into a sentiment classifier. More details of the proposed model are given as follows:

3.1 Contextual encoder with contrastive learning

The architecture of the contrastive learning-based contextual encoder is shown in Figure 3. Let $X = [w_1, \dots, w_a, \dots, w_{a+m-1}, \dots, w_n]$ be a sentence of n words and $A = [w_a, \dots, w_{a+m-1}]$ be the aspect of m words within S . The contrastive learning scheme during sentence encoding is implemented via data augmentation, feature extraction, and contrastive loss construction. Inspired by the data augmentation in image recognition [26]; [27], positive samples of the same polarities are generated using synonym substitution and random noise injection. Specifically, synonym substitution refers to randomly replacing words within the sentence with their synonyms from WordNet, while noise injection indicates introducing more aspect words and neutral sentiment words to the sentence. The sentiment is enhanced in (1):

$$X^E = \text{Enhance}(X). \tag{1}$$

The original sentence X and the data-enhanced sentence X^E are mapped to word vectors within the same feature space. We use the BERT model obtained through large-scale corpus training by Kenton et al to enhance the semantics of word representations. We then train the BERT model in a fine-tuned manner by freezing part of its parameters, which is written in (2):

$$h^{CLS}, H^X, H^A = \text{BERT}(CLS, X, SEP, A, SEP), \tag{2}$$

where CLS and SEP are BERT tokens representing the overall representation and the separation of the sentence, respectively. We thus obtain the sentence-level feature representation h^{CLS} , the word-level feature representation H^X , and the aspect feature representation H^A . Assuming that a batch consisting of k sentences is the model input for training, the sentence set composed of the original and the enhanced sentences is $X^{all} = [X^{batch}, (X^E)^{batch}] = [X_1, X_2, \dots, X_{2k}]$, with the corresponding sentiment polarity set denoted as $Y^{all} = [Y_1, Y_2, \dots, Y_{2k}]$. We also have the index set of all sentences as $I = [1, 2, \dots, 2k]$. For each sentence in X^{all} , a set of contrastive learning-based sentences with the same sentiment polarity is generated, i.e., $P^{all} = [P_1, P_2, \dots, P_{2k}]$, where $P_i = \{p: p \in I (Y_p = Y_i) \wedge (p \neq i)\}$. The contrastive learning loss of the contextual encoder is defined in (3):

$$\mathcal{L}_{ECL} = \sum_{i=1}^{2k} \frac{-1}{|P_i|} \sum_{p \in P_i} \log \frac{\exp(h_i^{CLS*} h_p^{CLS} / \tau)}{\sum_{k \in I/i} \exp(h_i^{CLS*} h_k^{CLS} / \tau)}, \tag{3}$$

where τ is a hyperparameter, indicating the temperature coefficient of contrastive learning. The higher the temperature coefficient is, the smaller the sum of the loss reaches. The parameter h_i^{CLS} stands for the representation of the i th sentence in X^{all} after BERT coding. In such a manner, the context representations of the same sentiment can be brought closer, and those of different sentiments are separated, improving the use of contextual and sentiment labels. Based on contrastive learning, abundant semantic information is integrated into the encoder, targeting at deriving context representation with key information.

3.2 Weighted label-enhanced syntactic GCN

The framework of the syntactic GCN via weighted label enhancement is presented in Figure 4. The syntax dependency of the input sentence is derived using the spaCy toolkit. Specifically, the sentence syntax dependency is characterized by a triplet, i.e., $(w_i, w_j, r_{i,j})$, where words w_i and w_j are of the relation

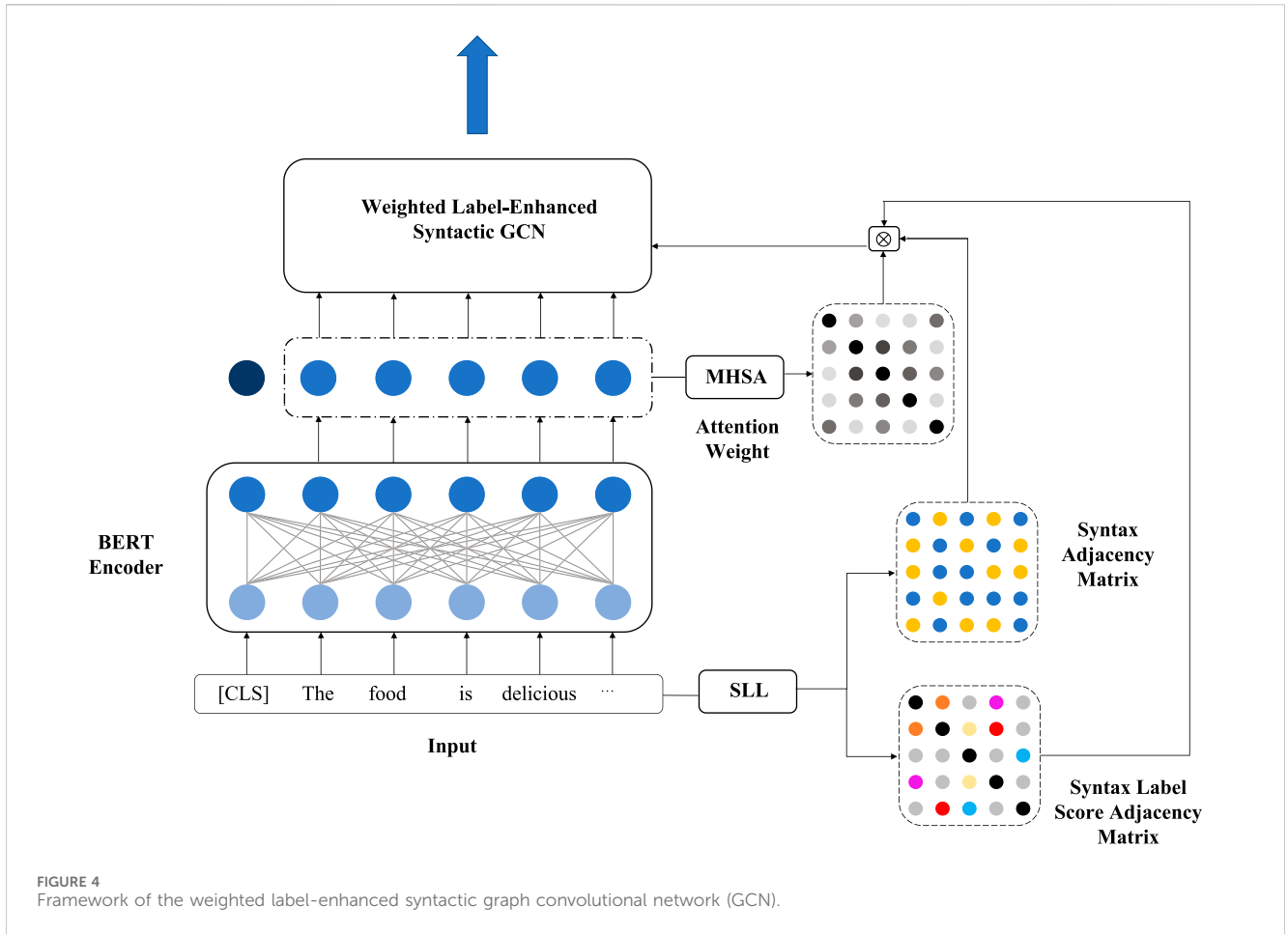


FIGURE 4 Framework of the weighted label-enhanced syntactic graph convolutional network (GCN).

r_{ij} . In line with the sentence syntax, we construct a syntax adjacency matrix $A^S \in \mathbb{R}^{n \times n}$ that denotes the connecting edges of the syntax dependency tree.

To address the effects of various syntactic labels in the sentiment classification, a syntax label learning (SLL) unit is built. The main purpose of the SLL unit is to transform the syntax label matrix to a learnable syntax label score matrix.

A lexicon $R = \{relation1: 1, relation2: 2, \dots, relation: t\}$ that consists of all syntactic relations from the corpus is constructed, from which each syntactic label is mapped to an index number. For each input sentence, all index numbers denote the syntactic relations consisting of a syntax label matrix. Then, the syntax label adjacency matrix $A^L \in \mathbb{R}^{n \times n}$ is built based on both the syntax dependency tree and the lexicon R . All syntax labels can be mapped into a unified feature space. The weighted score of each syntactic relation is thus resolved in (4), which is written as a syntax label score adjacency matrix $A^{LS} \in \mathbb{R}^{n \times n}$:

$$A^{LS} = Emb(A^L) * W_L * W_S, \quad (4)$$

where $Emb(\cdot)$ represents transforming the syntax label matrix into a learnable matrix for syntax label characterization, $W^L \in \mathbb{R}^{d_L \times d_S}$ and $W^S \in \mathbb{R}^{d_S \times 1}$ are learnable parameter matrices, and d_L and d_S are dimensions of A^L and A^S , respectively. We also have $Emb(A^L) \in \mathbb{R}^{n \times n \times d_{LS}}$, with d_{LS} standing for the dimension of the syntax label score space.

Likewise, the same syntactic relation type can have different degrees of importance within different semantic contexts. For this reason, the semantics among words are also integrated into the computation of the syntax label score. We take the multi-head self-attention (MHSA) mechanism to learn the semantic features and to revise the syntax label scores based on attentive weights. Notably, the elements in A^{LS} represent all the syntactic relation scores, which are not zero. To preserve the original syntax dependencies and remove irrelevant syntactic information, the basic syntax adjacency matrix is also used. The weighted syntax label adjacency matrix can be computed in (5):

$$A_{ij}^{WL} = A_{ij}^S * A_{ij}^W * A_{ij}^{LS} \quad (5)$$

with (6, 7, and 8)

$$A^W = Norm(Concat(head_1, head_2, \dots, head_h) \cdot W^{head}), \quad (6)$$

$$head_p = \frac{QW_p^Q \times (KW_p^K)^T}{\sqrt{d_{head}}}, \quad (7)$$

$$Q = W = H^X, \quad (8)$$

where A_{ij}^S is the syntax adjacency matrix of w_i and w_j parsed from the syntax dependency tree; A_{ij}^{LS} is the syntax label score adjacency matrix; A_{ij}^W stands for the semantic weight adjacency matrix derived from the MHSA mechanism; $Concat$ represents the vector

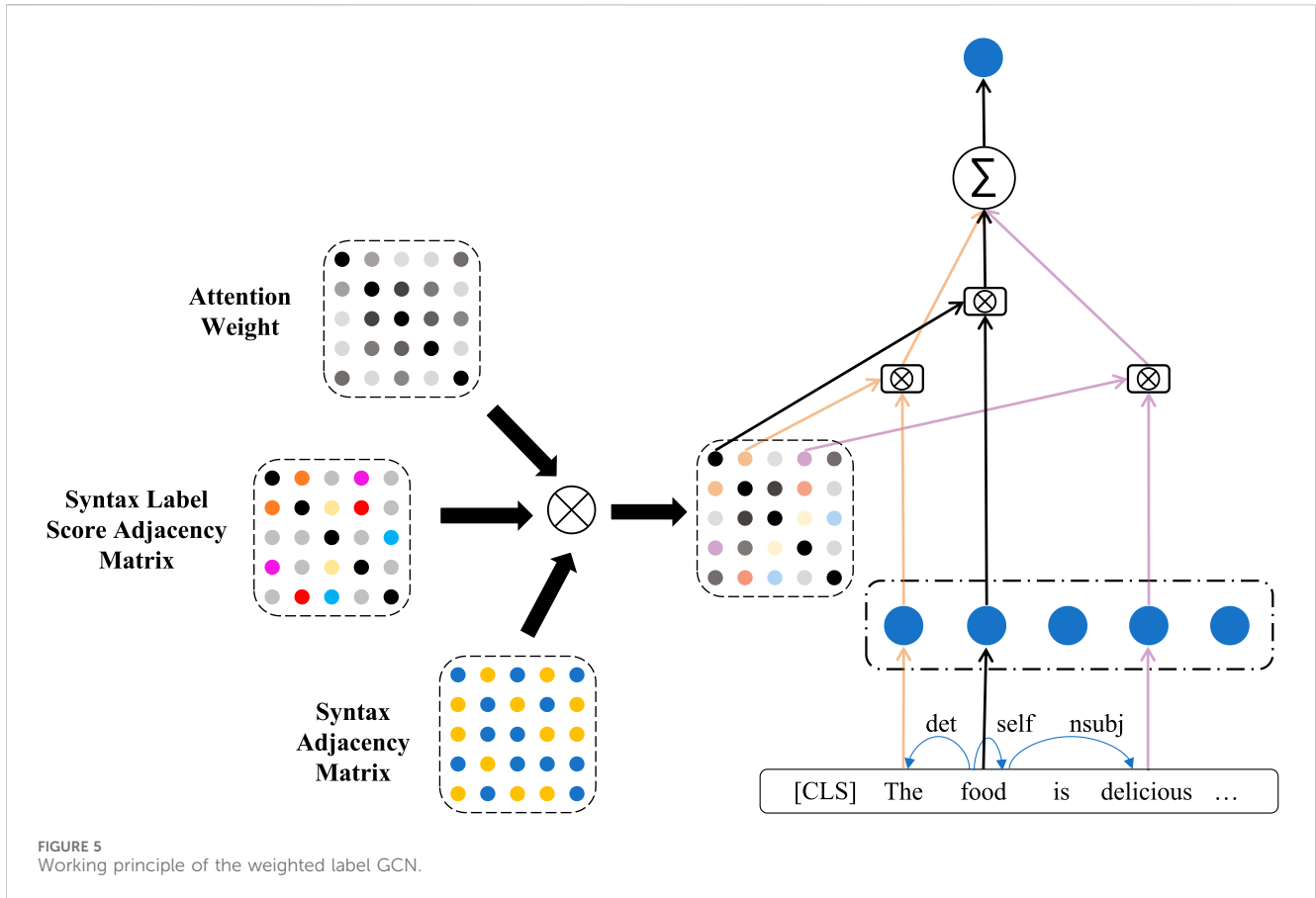


FIGURE 5 Working principle of the weighted label GCN.

concatenation; W^{head} is the parameter matrix during concatenation; $Norm(\cdot)$ is the normalization operation on the attentive weight matrix; W_p^Q and W_p^K are parameter matrices of the p th attention head in MHSA; d_{head} denotes the vector dimensions of each head; and h is a hyperparameter indicating the attention head number.

The working principle of the weighted label-enhanced syntactic GCN is shown in Figure 5. The input of the GCN is the weighted syntax label adjacency matrix A^{WL} and the feature representation H^X from BERT. The learning of syntactic information is derived in (9):

$$h_i^l = \sigma \left(\sum_{j=1}^n A_{ij}^{WL} W_s^l h_j^l + b_s^l \right), \quad (9)$$

where $H^0 = H^X = [h_1^X, h_2^X, \dots, h_n^X]$, h_j^l refers to the word vector of the j th word in the l th layer of the GCN, with l as an integer and $l \in [0, F]$, F is the layer number of the GCN, W_s^l is the learnable parametric matrix of the l th layer, b_s^l is the bias vector, and σ is an activation function. The output of the weighted label-enhanced syntactic GCN is the output of the last layer, i.e., $H^{out} = [h_1^F, h_2^F, \dots, h_n^F]$.

3.3 Syntax label-based contrastive learning scheme

Considering the effect of syntactic information in ABSC, the node pairs with the same syntax label indicate similar syntactic

features, and those with different syntax labels have differentiated features. As such, a contrastive learning scheme using syntax labels is proposed, aiming to enhance the learning of syntactic features at a higher level.

Assuming that K' triplets are of syntax dependencies within all the K sentences, the node-pair set of these triplets is $X' = [X'_1, X'_2, \dots, X'_{K'}]$. The syntax label set of these node pairs is $R' = [r_1, r_2, \dots, r_{K'}]$ with the index set $I' = [1, 2, \dots, K']$. Moreover, for each node pair in X' , a set of node pairs with the same syntax label for contrastive learning is constructed, i.e., $P_{m'}^l = \{p': p' \in I', (r_{p'} = r_{p'}) \wedge (p' \neq m')\}$. The syntax label-based contrastive learning loss is defined in (10):

$$\mathcal{L}_{LCL} = \sum_{m'=1}^{K'} \frac{-1}{|P_{m'}^l|} \sum_{p' \in P_{m'}^l} \log \frac{\exp(g_{m'}^* g_{p'} / \tau')}{\sum_{t \in I' / m'} \exp(g_{m'}^* g_t / \tau')}, \quad (10)$$

together with (11)

$$g_{m'} = (X'_{m'}[1]W_1^d + X'_{m'}[2]W_2^d)W_3^d + b^d, \quad (11)$$

where τ' is the temperature coefficient for contrastive learning and $g_{m'}$ represents the semantic feature representation by mapping the node-pair representations from the syntax dependency triplet and is normalized before the contrastive learning loss computation. We define $X'_{m'}[1]$ as the feature representation of the first node in the m' th node pair in X' and $X'_{m'}[2]$ as the feature representation of the second node in the m' th node pair. Both $X'_{m'}[1]$ and $X'_{m'}[2]$ are obtained from the BERT encoder, which convey semantic information. In addition,

W_1^{cl} , W_2^{cl} , and W_3^{cl} are learnable parameter matrices, and b^cl is a bias vector.

3.4 Feature fusion

Average pooling is performed on H^{out} to obtain the syntactic information-enhanced feature representation H^{out} in (12), which is further concatenated with h^{CLS} derived from the BERT encoder. The final sentence representation \tilde{H} is given in (13).

$$h^{out} = avgpool(h^{out}), \tag{12}$$

$$\tilde{H} = h^{out} \oplus h^{CLS}, \tag{13}$$

where \oplus denotes the concatenation operation. The final sentence representation \tilde{H} is sent to a Softmax classifier to obtain the sentiment polarity in (14):

$$y = softmax(W_o \tilde{H} + b_o). \tag{14}$$

The pseudocode of the proposed model is given as follows:

```

1: Input: data  $D$ , batch_size  $N$ .
2: Output: Sentiment polarity  $y$ 
3: for  $i = 0$  to  $n$  by  $N$  do
4:    $batch \leftarrow D [i : i + N]$ 
5:   for  $j$  in  $[i, i + batch\_size)$  do
6:      $h_j^{CLS}, H_j^X = BERT(X_j)$ 
7:      $A_S, A_W \leftarrow SLL(X_j)$ 
8:      $A^W \leftarrow MHSA(H_j^X)$ 
9:      $A^{WL} = A^S * A^W * A^{LS}$ 
10:     $H_j^{out} \leftarrow Weighted\_Label\_Enhanced\_GCN(A^{WL}, H_j^X)$ 
11:     $\tilde{H}_j \leftarrow Concatenate\_Features(H_j^{out}, h_j^{CLS})$ 
12:     $y_j = Softmax(\tilde{H}_j)$ 
13:   end for
14:    $\mathcal{L}_{total} = \mathcal{L}_{CE} + \alpha \mathcal{L}_{ECL} + (1 - \alpha) \mathcal{L}_{LCL}$ 
15:   Update network by combined loss  $\mathcal{L}_{total}$ 
16: end for

```

Algorithm 1. Dual contrastive learning-based GCN forward propagation algorithm.

3.5 Model training

Model training is implemented using cross-entropy and regularization as the loss function in (15):

$$\mathcal{L}_{CE} = - \sum_{(x,a) \in D} \sum_{c \in C} y_{(x,a)}^c \log \hat{y}_{(x,a)}^c + \lambda \|\theta\|^2, \tag{15}$$

where (x, a) represents the vector of a sentence–aspect pair; C refers to the set of sentiment classes; $y_{(x,a)}^c$ is the ground-truth sentiment distribution of (x, a) with sentiment C , and $\hat{y}_{(x,a)}^c$ is the predicted one; and λ is the coefficient of \mathcal{L}_2 regularization.

On account of the training of contrastive learning in our model, the total loss function \mathcal{L}_{total} is composed of the contrastive learning loss \mathcal{L}_{ECL} from the contextual encoder, and the

TABLE 1 Statistics of datasets.

Dataset	Positive		Neutral		Negative	
	Train	Test	Train	Test	Train	Test
Twitter	1,561	173	3,127	346	1,560	173
Laptop 14	994	341	464	169	870	128
Restaurant 14	2,164	728	637	196	807	196
Restaurant 15	912	326	36	34	256	182
Restaurant 16	1,240	469	69	30	439	117

contrastive learning loss \mathcal{L}_{LCL} based on the syntax label and the cross-entropy loss \mathcal{L}_{CE} is shown in (16):

$$\mathcal{L}_{total} = \mathcal{L}_{CE} + \alpha \mathcal{L}_{ECL} + (1 - \alpha) \mathcal{L}_{LCL}, \tag{16}$$

where α is a learnable coefficient to adjust the weights of contrastive learning losses in loss function.

4 Experiment

4.1 Experimental setup

The working performance of the DCL-GCN is evaluated on five benchmark datasets, which are Restaurant 14, Restaurant 15, Restaurant 16, and Laptop 14 from SemEval [28]; [29,30], and Twitter [31]. The sentiment of each aspect from the datasets is labeled as positive, neutral, or negative.

Following the idea of [15], the sentences labeled as conflicting sentiment or without explicit aspects from Restaurant 15 and Restaurant 16 are removed. Details of each dataset are given in Table 1.

In this experiment, the lexicon size of the BERT model is set to 30,522, the word embedding dimension is 768, and the layer number of the transformer is 12. The head number of the MHSA is 8, and the learning rate is 0.00001. The layer number of the weighted label-enhanced syntactic GCN is 2. Both τ and τ' in contrastive learning schemes are set to 0.02. The \mathcal{L}_2 regularization coefficient is 0.00001. An Adam optimizer is adopted during training with a data batch size of 32. All the hyperparameters used in the experiment are given in Table 2.

4.2 Baseline

In order to verify the effectiveness of the DCL-GCN in ABSC, five state-of-the-art methods are taken for comparison:

- **BERT [32]:** The basic BERT model is established based on the bidirectional transformer. With the concatenation of sentences and the corresponding aspect, BERT can be applied to ABSC.
- **BERT4GCN [33]:** The BERT model and GCN are integrated, which exploits sequential features and positional information to augment the model learning.
- **R-GAT + BERT [25]:** The pre-trained BERT is integrated with the R-GAT, where BERT is used for sentence encoding.

TABLE 2 Parameter settings.

Parameter	Value
BERT model lexicon size	30,522
Word embedding dimension	768
Transformer layers	12
Multi-head self-attention (MHSA) heads	8
Learning rate	0.00001
Weighted label-enhanced syntactic graph convolutional network (GCN) layers	2
T	0.02
r'	0.02
\mathcal{L}_2 regularization coefficient	0.00001
Batch size	32
Optimizer	Adam

- **DGEDT + BERT** [34]: The pre-trained BERT is integrated with DGEDT, where BERT is used for sentence encoding.
- **TGCN + BERT** [35]: The dependency type is identified with type-aware graph convolutional networks, while the relation is distinguished with an attention mechanism. The pre-trained BERT is used for sentence encoding.

All results are expressed in percentage values. “-” denotes that the results are not reported in the published research article. The best performance achieved is marked in bold.

4.3 Result analysis

We take two metrics, i.e., accuracy and Macro-F1, to evaluate the working performance of the proposed model. Table 3 shows the results of six different methods on the task of ABSC. One can observe that our model achieves the best and most consistent result among all the evaluation settings. It is clear that the DCL-GCN result is more remarkable than a range of competitive baselines on all five benchmark datasets. In line with these results, the following observations are made.

First, our model achieves the best and most consistent result among all the evaluation settings. The minimum performance gaps between the DCL-GCN and the baselines are 1.33% (against the R-GAT) on Restaurant 14, 1.62% (against the T-GCN) on Restaurant 15, 1.33% (against the T-GCN) on Restaurant 16, 1.54% (against DGEDT) on Laptop 14, and 0.22 (against DGEDT) on Twitter. In addition, the F1 values on Restaurant 15 and Restaurant 16 are 3.64% (against the T-GCN) and 5.04% (against DGEDT), respectively, higher than the best-performing baseline method, which are significant.

Second, the syntax-dependent-method (BERT4GCN) performs worse than models integrated with both syntax dependency and syntactic relations (R-GCT and T-GCN). The main reason is that the deeper-level syntactic information can be neglected by solely exploiting the dependencies among words. By contrast, the syntactic relation encoded in our model benefits the sentiment comprehending to a large extent. The highest accuracy of our

model reaches 93.65 on Restaurant 16, indicating the importance of syntax dependency and syntactic relations in ABSC.

Third, compared with other baselines, the basic BERT model has its own distinctiveness in tackling sentence semantic information. By incorporating BERT into state-of-the-art methods, the working performance is substantially improved, which is the outcome of our model. Notably, the proposed model significantly outperforms the baselines, demonstrating that the contextual semantics take full advantage in line with the BERT-based contrastive learning scheme.

It is worth noting that the DCL-GCN gives rise to the enhancement in both syntax and semantics learning. With the application of the dual contrastive learning scheme, it is reasonable to expect better working performance in ABSC, as it is the case.

4.4 Ablation study

The impact of different components in our model is investigated by conducting an ablation study (Table 4). *w/o L_{ECL}* specifies that the contrastive learning scheme of the contextual encoder is removed; *w/o L_{LCL}* specifies that the syntax label-based contrastive learning scheme is removed; and *w/o WL-GCN* indicates that the weighted label-enhanced syntactic GCN is ablated.

As presented in Table 4, the most significant module in our model is the weighted label-enhanced syntactic GCN. The exploiting of syntactic information shows its effectiveness in word sentiment learning. With the sole utilization of semantics, even with a contrastive learning strategy, the working performance is inferior to the syntactic-based methods in all evaluation settings. Clearly, the integration of semantics and syntax has superiority in ABSC tasks. Moreover, the removal of the contrastive learning scheme from the contextual encoder leads to a substantial decrease on all five datasets. The performance decreases of the accuracy and F1 score on Twitter are 2.76% and 2.25%, respectively. As a result, the contrastive learning scheme in the BERT encoder effectively promotes semantic information learning. By contrast, the syntax label-based contrastive learning scheme makes a relatively small contribution to the model. We can infer that the application of

TABLE 3 Experimental results on five public datasets.

Model	Twitter		Laptop 14		Restaurant 14		Restaurant 15		Restaurant 16	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
BERT [32]	75.00	72.53	78.68	74.64	84.55	77.34	83.40	65.28	89.54	70.47
BERT4GCN [33]	74.73	73.76	77.49	73.01	84.75	77.11	-	-	-	-
R-GAT + BERT [25]	76.15	74.88	78.21	74.07	86.60	81.35	-	-	-	-
DGEDT + BERT [34]	77.90	75.40	79.80	75.60	86.30	80.00	84.00	71.00	91.90	79.00
TGCN + BERT [35]	76.45	75.25	80.88	77.03	86.16	79.95	85.26	71.69	92.32	77.29
Our DCL-GCN + BERT	78.12	76.37	82.42	79.20	87.93	82.53	86.88	75.35	93.65	84.04

The bold values represent the best performance achieved among the different models or methods compared in the table. Specifically, the bold values indicate the highest accuracy and Macro-F1 scores obtained for each dataset (Twitter, Laptop14, Restaurant14, Restaurant15, Restaurant16) in the aspect-based sentiment classification (ABSC) task. These bold values highlight the superior results of the model we proposed compared to the baseline methods, showcasing its effectiveness in sentiment classification across different datasets.

TABLE 4 Results of the ablation study.

Model	Twitter		Laptop 14		Restaurant 14		Restaurant 15		Restaurant 16	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
<i>w/o</i> L_{ECL}	75.36	74.12	81.28	77.49	85.94	80.18	84.63	72.78	91.44	81.64
<i>w/o</i> L_{LCL}	76.12	74.89	81.13	77.30	86.12	80.67	85.24	73.68	92.17	82.26
<i>w/o</i> WL-GCN	75.13	73.85	80.62	76.83	85.20	79.47	84.27	72.14	91.23	81.08
Full model	78.12	76.37	82.42	79.20	87.93	82.53	86.88	75.35	93.65	84.04

The bold values represent the best performance achieved among the different models or variations compared in the table. Specifically, the bold values indicate the highest accuracy and Macro-F1 scores obtained for each dataset (Twitter, Laptop14, Restaurant14, Restaurant15, Restaurant16) in the aspect-based sentiment classification (ABSC) task when specific components or modules of the proposed model are included.

syntax labels also enhances the use of syntactic information and, thus, contributes to the sentiment classification.

significantly increases, resulting in overfitting issues during training. In this way, a test accuracy decrease is inevitable.

4.5 Impact of hyperparameters

An experiment is carried out to analyze the effect of the self-attention head number on model working performance. The head number of the self-attention network is set to [1, 2, 3, . . . , 8]. The model accuracy with different head numbers is presented in Figure 6.

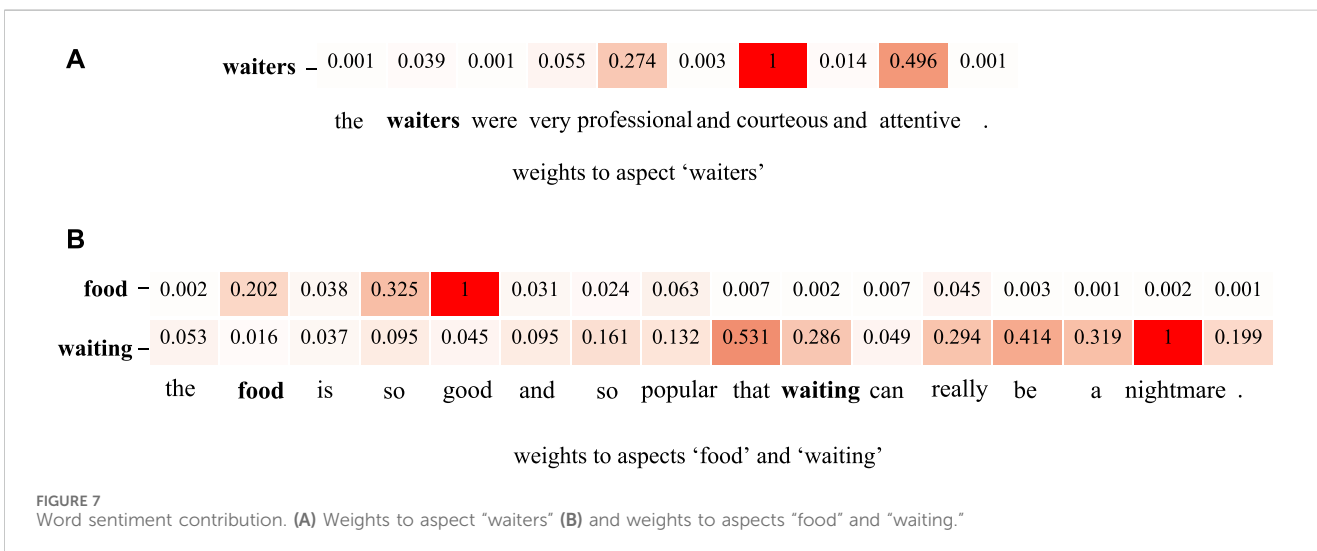
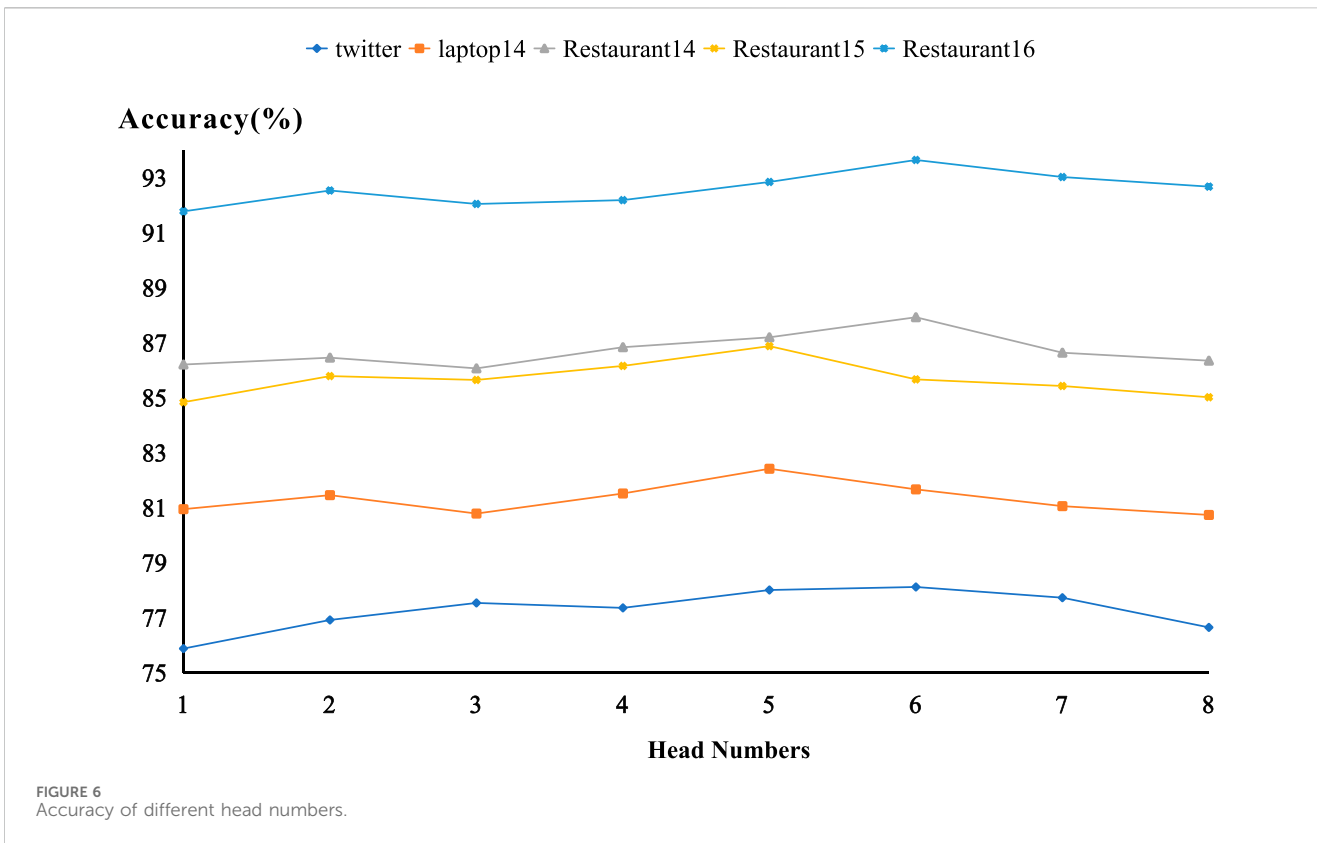
Apparently, the DCL-GCN achieves the highest accuracy with a head number of 5 on Laptop 14 and Restaurant 15 and a head number of 6 on Twitter, Restaurant 14, and Restaurant 16. In line with the multi-head self-attention mechanism, the attention head stands for the vector representation in feature spaces via different mapping methods. When the number of attention heads is reduced, the self-attention mechanism operates within a smaller space with correspondingly fewer semantic features. Accordingly, the proposed model fails to capture sufficient semantic information. On the other hand, when the head number exceeds 6, the model parameter size

4.6 Case study

Two samples are selected to visualize the working performance, in order to further validate the distinctiveness of DCL-GCN. Specifically, the representations of the sentence and the words are maintained. We shall define a parameter φ as the contribution of each word for sentiment delivery in the sentence, which is defined in Eq. 17:

$$\varphi(X, w_i) = \frac{|\tilde{H} - \tilde{H}_{w_i}|}{\sum_{j=1}^n |\tilde{H} - \tilde{H}_{w_j}|} \tag{17}$$

The sentiment contribution of each word is shown in Figure 7. For the sample given in Figure 7A, the contextual words “professional,” “courteous,” and “attentive” make the largest contribution toward the aspect “waiters.” Our model is capable of extracting the most informative words for sentiment



expressing. The sentence in Figure 7B contains two aspects, i.e., “food” and “waiting.” For the aspect word “food,” the proposed model accurately identifies the top two highest sentiment contribution words as “good” and “so.” Regarding “waiting,” not only is the the sentiment word “nightmare” captured but also the syntactic relation words “so...that...” for resultative adverbial clause establishment. Both semantics and syntax are used for sentiment classification.

In our model, the use of contrastive learning enhances the learning of sentence semantics, and the build of the weighted

label-enhanced syntactic GCN fully exploits the syntactic information. The integration of semantic information and syntactic information leads to a competitive manner in ABSC.

5 Conclusion

In this work, we propose a GCN based on dual contrastive learning and syntax label enhancement for ABSC tasks. To obtain sentiment information, a contrastive learning scheme is integrated

to a BERT encoder to enhance the learning of semantic-related contextual information. Then, our model exploits both the syntax dependency and syntactic relation, based on which a weighted label-enhanced syntactic GCN is established. In addition, the learning of the syntax label is enhanced using contrastive learning. A syntactic triplet between words is mapped into a unified feature space for syntax and semantic integration. The experimental results reveal that the proposed model achieves state-of-the-art performance on five benchmark datasets. The ablation study, the hyperparameter analysis experiment, and the case study also obtain superior working performance.

Future work will focus on introducing more information for further improving the accuracy of ABSC and other sentiment analysis tasks, such as background knowledge and part-of-speech information. In addition, the integration of different categories of information into the model is also considered.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

YH: conceptualization, methodology, and writing—original draft. AD: conceptualization, methodology, and writing—original draft. SC: formal analysis, methodology, and writing—original draft. QK: conceptualization, formal analysis, and writing—original draft. HZ: funding acquisition, supervision, and

writing—review and editing. QC: supervision and writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported in part by the Guangdong Basic and Applied Basic Research Foundation (no. 2023A1515011370), the National Natural Science Foundation of China (no. 32371114), and the Characteristic Innovation Projects of Guangdong Colleges and Universities (no. 2018KTSCX049).

Conflict of interest

Author YH was employed by Datastory, author AD was employed by China Merchants Bank, and author QK was employed by Guangzhou Qizhi Information Technology Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Pang B, Lee L. Opinion mining and sentiment analysis. *Foundations Trends® in information retrieval*. Now Publishers Inc. (2008) 2 (1–2):1–135. doi:10.1561/1500000011
- Bing L. *Sentiment analysis and opinion mining (synthesis lectures on human language technologies)*. Chicago, IL, USA: University of Illinois (2012).
- Zheng Y, Zhang R, Mensah S, Mao Y. Replicate, walk, and stop on syntax: an effective neural network model for aspect-level sentiment classification. *Proc AAAI Conf Artif Intelligence* (2020) 34:9685–92. doi:10.1609/aaai.v34i05.6517
- Tsytarou M, Palpanas T. Survey on mining subjective data on the web. *Data Mining Knowledge Discov* (2012) 24:478–514. doi:10.1007/s10618-011-0238-6
- Lin T, Joe I. An adaptive masked attention mechanism to act on the local text in a global context for aspect-based sentiment analysis. *IEEE Access* (2023) 11:43055–66. doi:10.1109/access.2023.3270927
- Žunić A, Corcoran P, Spasić I. Aspect-based sentiment analysis with graph convolution over syntactic dependencies. *Artif Intelligence Med* (2021) 119:102138. doi:10.1016/j.artmed.2021.102138
- Yusuf KK, Ogbuju E, Abiodun T, Oladipo F. A technical review of the state-of-the-art methods in aspect-based sentiment analysis. *J Comput Theories Appl* (2024) 2:67–78. doi:10.62411/jcta.9999
- He Y, Huang X, Zou S, Zhang C. Psan: prompt semantic augmented network for aspect-based sentiment analysis. *Expert Syst Appl* (2024) 238:121632. doi:10.1016/j.eswa.2023.121632
- Huang X, Li J, Wu J, Chang J, Liu D, Zhu K. Flexibly utilizing syntactic knowledge in aspect-based sentiment analysis. *Inf Process Manag* (2024) 61:103630. doi:10.1016/j.ipm.2023.103630
- Wang P, Tao L, Tang M, Wang L, Xu Y, Zhao M. Incorporating syntax and semantics with dual graph neural networks for aspect-level sentiment analysis. *Eng Appl Artif Intelligence* (2024) 133:108101. doi:10.1016/j.engappai.2024.108101
- Nazir A, Rao Y, Wu L, Sun L. Iaf-ig: an interactive attention fusion network with local and global perspective for aspect-based sentiment analysis. *IEEE Trans Affective Comput* (2022) 13:1730–42. doi:10.1109/taffc.2022.3208216
- Gou J, Sun L, Yu B, Wan S, Ou W, Yi Z. Multilevel attention-based sample correlations for knowledge distillation. *IEEE Trans Ind Inform* (2022) 19:7099–109. doi:10.1109/tii.2022.3209672
- Tang D, Qin B, Feng X, Liu T. Effective lstms for target-dependent sentiment classification (2015). Available at: <https://arxiv.org/abs/1512.01100> (Accessed December 3, 2015).
- Wang Y, Huang M, Zhu X, Zhao L. Attention-based lstm for aspect-level sentiment classification. In: Proceedings of the 2016 conference on empirical methods in natural language processing; November, 2016; Austin, Texas (2016). p. 606–15.
- Chen P, Sun Z, Bing L, Yang W. Recurrent attention network on memory for aspect sentiment analysis. In: Proceedings of the 2017 conference on empirical methods in natural language processing; September, 2017; Copenhagen, Denmark (2017). p. 452–61.
- Sun C, Huang L, Qiu X. Utilizing bert for aspect-based sentiment analysis via constructing auxiliary sentence (2019). Available at: <https://arxiv.org/abs/1903.09588> (Accessed March 22, 2019).
- Gou J, Yuan X, Yu B, Yu J, Yi Z. Intra-and inter-class induced discriminative deep dictionary learning for visual recognition. *IEEE Trans Multimedia* (2023) 25:1575–83. doi:10.1109/tmm.2023.3258141
- Gou J, Xie N, Liu J, Yu B, Ou W, Yi Z, et al. Hierarchical graph augmented stacked autoencoders for multi-view representation learning. *Inf Fusion* (2024) 102:102068. doi:10.1016/j.inffus.2023.102068

19. Suresh V, Ong DC. Not all negatives are equal: label-aware contrastive loss for fine-grained text classification (2021). Available at: <https://arxiv.org/abs/2109.05427> (Accessed September 12, 2021).
20. Wang S-H, Govindaraj VV, Górriz JM, Zhang X, Zhang Y-D. Covid-19 classification by fgcn with deep feature fusion from graph convolutional network and convolutional neural network. *Inf Fusion* (2021) 67:208–29. doi:10.1016/j.inffus.2020.10.004
21. Zhang Y-D, Satapathy SC, Guttery DS, Górriz JM, Wang S-H. Improved breast cancer classification through combining graph convolutional network and convolutional neural network. *Inf Process Manag* (2021) 58:102439. doi:10.1016/j.ipm.2020.102439
22. Sun K, Zhang R, Mensah S, Mao Y, Liu X. 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); November, 2019; Hong Kong, China (2019). p. 5679–88.
23. Zhang C, Li Q, Song D. Aspect-based sentiment classification with aspect-specific graph convolutional networks (2019). Available at: <https://arxiv.org/abs/1909.03477> (Accessed September 8, 2019).
24. Liang B, Yin R, Gui L, Du J, Xu R. 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; December, 2020; Barcelona, Spain (Online) (2020). p. 150–61.
25. Wang K, Shen W, Yang Y, Quan X, Wang R. Relational graph attention network for aspect-based sentiment analysis (2020). Available at: <https://arxiv.org/abs/2004.12362> (Accessed April 26, 2020).
26. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: International conference on machine learning (PMLR); July, 2020; Virtual (2020). p. 1597–607.
27. Khosla P, Teterwak P, Wang C, Sarna A, Tian Y, Isola P, et al. Supervised contrastive learning. *Adv Neural Inf Process Syst* (2020) 33:18661–73. doi:10.48550/arXiv.2004.11362
28. Kirange D, Deshmukh RR, Kirange M. Aspect based sentiment analysis semeval- 2014 task 4. *Asian J Comp Sci Inf Tech (Ajsit)* (2014) 4. doi:10.15520/ajcsit.v4i8.9
29. Pontiki M, Galanis D, Papageorgiou H, Manandhar S, Androutsopoulos I. Semeval-2015 task 12: aspect based sentiment analysis. In: Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015); June, 2015; Denver, Colorado (2015). p. 486–95.
30. Pontiki M, Galanis D, Papageorgiou H, Androutsopoulos I, Manandhar S, Al-Smadi M, et al. Semeval-2016 task 5: aspect based sentiment analysis. In: ProWorkshop on Semantic Evaluation (SemEval-2016) (Association for Computational Linguistics); June, 2016; San Diego, California, USA (2016). p. 19–30.
31. Dong L, Wei F, Tan C, Tang D, Zhou M, Xu K. Adaptive recursive neural network for target-dependent twitter sentiment classification. In: Proceedings of the 52nd annual meeting of the association for computational linguistics; June, 2014; Baltimore, Maryland (2014). p. 49–54.
32. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: pre-training of deep bidirectional transformers for language understanding (2018). Available at: <https://arxiv.org/abs/1810.04805> (Accessed October 11, 2018).
33. Xiao Z, Wu J, Chen Q, Deng C. Bert4gcn: using bert intermediate layers to augment gcn for aspect-based sentiment classification (2021). Available at: <https://arxiv.org/abs/2110.00171> (Accessed October 1, 2021).
34. Tang H, Ji D, Li C, Zhou Q. Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification. In: Proceedings of the 58th annual meeting of the association for computational linguistics; July, 2020; Online (2020). p. 6578–88.
35. Tian Y, Chen G, Song Y. Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble. In: Proceedings of the 2021 conference of the North American chapter of the association for computational linguistics: human language technologies; June, 2021 (2021). p. 2910–22.