Knowledge-Based Systems 235 (2022) 107643

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

# Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks

Bin Liang [a,b], Hang Su [a,b], Lin Gui [c], Erik Cambria [d], Ruifeng Xu [a,b,e,*]

[a] School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China
[b] Joint Lab of HITSZ and China Merchants Securities, Shenzhen, China
[c] Department of Computer Science, University of Warwick, Coventry, UK
[d] School of Computer Science and Engineering, Nanyang Technological University, Singapore
[e] Peng Cheng Laboratory, Shenzhen, China

ARTICLE INFO

*Article history:*
Received 4 November 2020
Received in revised form 29 May 2021
Accepted 1 June 2021
Available online 29 October 2021

*Keywords:*
Sentiment analysis
Aspect sentiment analysis
Affective knowledge
Graph convolutional networks

ABSTRACT

Aspect-based sentiment analysis is a fine-grained sentiment analysis task, which needs to detection the sentiment polarity towards a given aspect. Recently, graph neural models over the dependency tree are widely applied for aspect-based sentiment analysis. Most existing works, however, they generally focus on learning the dependency information from contextual words to aspect words based on the dependency tree of the sentence, which lacks the exploitation of contextual affective knowledge with regard to the specific aspect. In this paper, we propose a graph convolutional network based on SenticNet to leverage the affective dependencies of the sentence according to the specific aspect, called **Sentic GCN**. To be specific, we explore a novel solution to construct the graph neural networks via integrating the affective knowledge from SenticNet to enhance the dependency graphs of sentences. Based on it, both the dependencies of contextual words and aspect words and the affective information between opinion words and the aspect are considered by the novel affective enhanced graph model. Experimental results on multiple public benchmark datasets illustrate that our proposed model can beat state-of-the-art methods.



## 1. Introduction

Sentiment analysis and opinion mining [1] have become a hot research topic in natural language processing (NLP) with the rapid development of social media communities. Different from conventional sentiment analysis that predicting the overall sentiment of a given text, aspect-based sentiment analysis (ABSA) aims to detect the sentiment polarities (e.g., positive, negative or neutral) of different aspects in the same sentence [2,3]. For example, in sentence "*The food is good, but the service is terrible*". Two aspects are mentioned in the sentence: "*food*" and "*service*". The sentiment polarity is positive towards the aspect "*food*", but is negative towards the aspect "*service*". Generally, this task is formulated as aspect detection and sentiment analysis for a specific aspect. In this paper, we investigate the task of determining the sentiment polarity with regard to a specific aspect.

Most early works for ABSA exploited neural models to model the sentiment information from the context for a given aspect [4–7]. And further, attention based neural networks have become the promising approaches for focusing on a specific aspect relevant parts from the context of a sentence [8,9]. While promising results have been achieved by attention-based models, they generally lack the function of extracting the sentiment dependencies between contextual words and aspect words and ignore the significance of sentiment-related words in the context, which limit the performance of ABSA.

Intuitively, syntactical information can provide the relation information between a specific aspect and the sentiment expression in ABSA. Therefore, syntactic dependencies from contextual words to aspect words should be considered in predicting sentiment polarity for a specific aspect. To address this issue, some graph networks based on dependency tree are proposed to model the structure of a sentence through the dependency tree of the sentence [10,11]. Hence, these methods can draw syntactically relevant words to the specific aspect, and capture long-term syntactic dependencies via convolutional operation.

Most existing graph network-based models, however, only considered the syntactic dependencies of a sentence and ignore commonsense knowledge information [12] when constructing

* Corresponding author at: School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China.
*E-mail addresses:* bin.liang@stu.hit.edu.cn (B. Liang), suhang@stu.hit.edu.cn (H. Su), lin.gui@warwick.ac.uk (L. Gui), cambria@ntu.edu.sg (E. Cambria), xuruifeng@hit.edu.cn (R. Xu).

https://doi.org/10.1016/j.knosys.2021.107643
0950-7051/

the graph. Hence, we argue that incorporating affective commonsense knowledge into the graph networks can promote the model to extract the sentiment-related dependencies between contextual words and the specific aspect. To pursue this goal, we explore a novel solution to construct the graph over both the dependency tree and affective commonsense knowledge for each sentence. To be specific, we first construct an ordinary dependency graph for each sentence based on the dependency tree to capture the syntactical information of the sentence. Secondly, to leverage the sentiment dependencies between contextual words and aspect words, the external affective commonsense knowledge is incorporated into the graph generation. Based on the syntactic dependencies of the sentence and the affective information with respect to the specific aspect, each sentence can be modeled as an aspect-specific affective enhanced dependency graph. Afterward, the affective enhanced dependency graph is fed as input into the GCN-based model to capture the graph representations of the sentence.

The main contributions of our work can be summarized as follows:

- The task of ABSA is reconsidered so as to leverage both the syntactic dependencies of the sentence and the affective information with respect to the specific aspect.
- A novel solution of constructing graph for graph convolutional networks over dependency tree and affective commonsense knowledge is proposed to capture affective dependencies corresponding to specific aspects.
- Experimental results on four benchmark datasets demonstrate the effectiveness of our proposed model in ABSA task.

The rest of the paper is organized as follows: Section 2 presents related works in ABSA, graph convolutional network (GCN), and affective commonsense knowledge; Section 3 describes our proposed model in detail; Section 4 presents the experiment setup and evaluation results; finally, Section 5 concludes the paper and outlines future directions.

## 2. Related work

### 2.1. Aspect-based sentiment analysis

In ABSA, existing neural network-based approaches generally attempt to learn the key sentiment information from the context and determine the sentiment polarity according to the specific aspect or target [4,8,9,13–19]. Among them, Tang et al. [13] utilized two long short-term memory (LSTM) networks to extract the bidirectional sentiment information from the context for a given aspect, which can extract the relations between context and the aspect for learning the important parts of context to aspect-related sentiment detection. Majumder et al. [4] incorporated the information of related neighboring aspects into the sentiment learning of the aspect via using memory networks, which iteratively models the influence from the other aspects to generate accurate aspect representation. For the attention-based method, Chen et al. [14] explored a multiple attention-based memory network to learn the long distance sentiment features. Fan et al. [16] proposed a multiple fine-grained attention model to interactively learn the relations between aspect words and contextual words. For convolutional neural network, Xue and Li [17] propose a gated convolutional neural network model to selectively output the sentiment features according to the given aspect based on gating units. For multi-task based methods, He et al. [15] consider both document- and aspect-level information in ABSA, and transfer the features learned from document-level to improve the sentiment learning of aspect-level. Further, He et al. [18] utilize an iterative message passing scheme to explicitly model the interactions between different tasks. Chen and Qian [20] explored a relation-aware collaborative learning framework to fully exploit the interactive relations of different subtasks in the complete ABSA task. For BERT-based method, Phan and Ogunbona [7] explored the grammatical aspect of the sentence and employs the self-attention mechanism for syntactical learning to enhance the performance of the aspect extractor for ABSA based on BERT. Besides, for capsule network-based models [21], Du et al. [22] construct vector-based feature representations and cluster features for aspect-related sentiment extraction based on an EM routing algorithm, Chen and Qian [23] transfer document-level knowledge to aspect sentiment classification with capsule network-based model.

### 2.2. Graph convolutional network

Recently, graph network-based models have achieve promising performance in many NLP tasks, such as text classification [24,25], relation extraction [26], Chinese NER [27,28]. Subsequently, graph neural networks have shown high potentials on ABSA. Zhang et al. [11] proposed a GCN over the dependency tree of the sentence to exploit syntactical information of the context and word dependencies for the aspect. Huang et al. [29] proposed a target-dependent graph attention network (GAT) to learn sentiment information for an aspect by exploring the dependency relations among contextual words. Sun et al. [10] stacked a GCN layer over an LSTM to the task of ABSA, which utilized a Bi-directional LSTM model to learn the contextual features of a sentence, and further performed convolutions over a dependency tree to extract rich representations for ABSA. Tang et al. [30] jointly considered the learning of flat representations learned from Transformer and graph-based representations by a dependency graph enhanced dual-transformer network. Wang et al. [31] proposed an aspect-oriented tree structure by reshaping and pruning ordinary dependency trees to focus on the target aspects in the sentence. However promising, these graph network-based models neglect the affective information between contextual opinion words and the aspect words, which is capable of directly presenting the sentiment expression from the sentence for the specific aspect. Hence, how to enhance the feature representations extracted by the GCN via affective knowledge should be considered in the task of ABSA.

### 2.3. Affective commonsense knowledge

The significant role of external knowledge in many NLP tasks has been remarked on [32–37]. Analogously, external affective commonsense knowledge has been typically used as a source for enhancing the sentiment feature representations in the task of sentiment analysis [5,38–42]. SenticNet is a publicly available resource for opinion mining and sentiment analysis, which uses dimensionality reduction to infer the polarity of commonsense concepts and provide an affective value for each concept [43–48]. In SenticNet, affective values are very close to 1 for those strong positive concepts, while for strong negative concepts, affective values are close to $-1$. As a versatile affective knowledge base, SenticNet reveals remarkable performance for enhancing the sentiment representation learning [33,49]. In which, Xing et al. [50] demonstrated that SenticNet performed overwhelmingly in comparison with other sentiment lexicons. Based on SenticNet, Ma et al. [51] incorporated commonsense knowledge into the LSTM model to extract target-level and sentence-level sentiment features in targeted ABSA. Hence, based on the remarkable performance achieves by SenticNet, we exploit SenticNet 6 [48], which contains 200,000 concepts, as a commonsense knowledge base to embellish the graph and then enhance the sentiment representations extracted by the graph model.

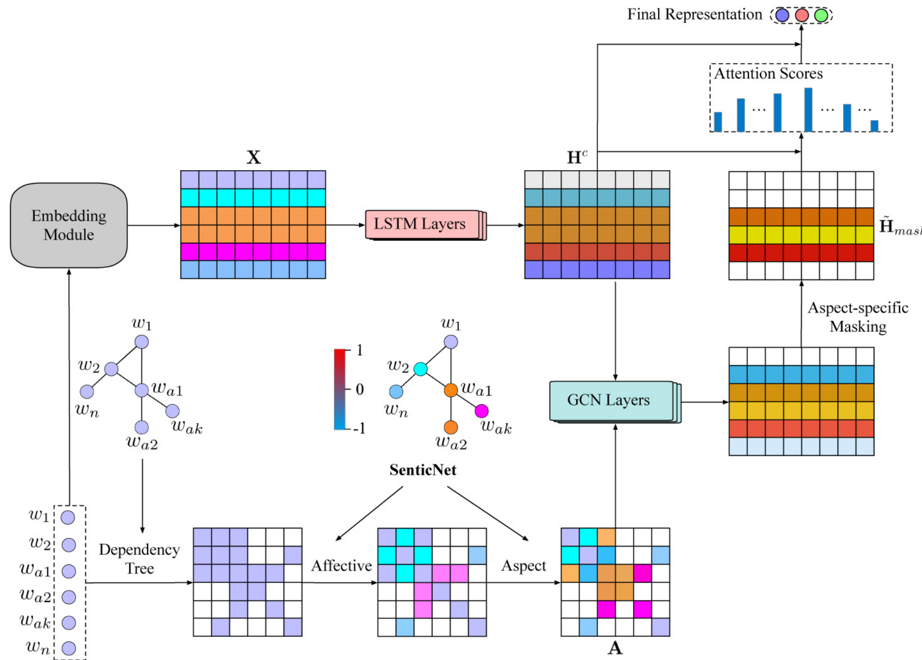



**Fig. 1.** Overall architecture of the proposed SenticNet-based graph convolutional network.

## 3. Proposed model

In this section, we illustrate the proposed SenticNet-based graph convolutional network (Sentic GCN) in detail. The overall architecture of the proposed model is demonstrated in Fig. 1.

### 3.1. Overview

As illustrated in Fig. 1, our proposed model consists of two primary components: (1) Learning contextual representations with LSTM layers, which takes the embedding matrix of each sentence as input and derives the hidden contextual representations of the input sentence. (2) Leveraging graph information with GCN layers, which takes the hidden contextual representations of the sentence and corresponding affective enhanced graph as input to capture the potential sentiment dependencies of the contextual words.[1] Afterward, the representations derived from these two components are combined to extract the significant sentiment dependencies with respect to the specific aspect. Here, different from most previous graph-based models, which only focus on the syntactical information of the sentence. To leverage the contextual sentiment dependencies with respect to the specific aspect, we emphasize the words with highly aspect-related sentiment coloring by explicably refining the graph structure of the sentence, since the aspect-related sentiment feature is the most important factor in ABSA task.

### 3.2. Task definition

Given a sentence consists of $n$ words: $s = \{w_1, w_2, \ldots, w_{a1}, w_{a2}, \ldots, w_{ak}, \ldots, w_n\}$, where $w_i$ represents the $i$th contextual word and $w_{ai}$ represents the $i$th aspect word. It should be noticed that a sentence may contain one or more aspects corresponding to different sentiment polarities, i.e., *Positive*, *Negative*, or *Neutral*. Here, each aspect may consist of single or multiple words. The goal of ABSA is to detect the sentiment polarity for a given aspect via extracting aspect-related sentiment information from the context.

[1] Here, the GCN-layers structure used in our work is inspired by a previous remarkable GCN-based study [11].

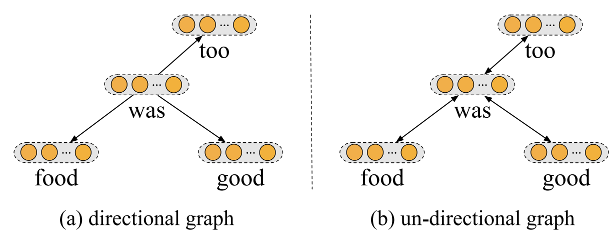



**Fig. 2.** A sketch of constructing directional and un-directional dependency graph for sentence "*food was good too*".

### 3.3. Embedding module

For each input sentence, we embed each word in the sentence into a $m$-dimensional embedding from the embedding lookup table $\mathbf{X} \in \mathbb{R}^{m \times |N|}$, and then we get the input embedding matrix $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{a1}, \mathbf{x}_{a2}, \ldots, \mathbf{x}_{ak}, \ldots, \mathbf{x}_n]$ of the input sentence. Where $\mathbf{x}_i \in \mathbb{R}^m$ is the word embedding of contextual word $w_i$ and where $\mathbf{x}_{ai} \in \mathbb{R}^m$ is the word embedding of aspect word $w_{ai}$, $m$ is the dimension of word vectors, $n$ is the length of sentence, and $|N|$ is the vocabulary size. The embedding lookup table is usually derived from pre-trained embeddings such as GloVe [52] or BERT [53] et al. and they are fine-tuned during the training process.

### 3.4. Constructing graph over dependency tree

To leverage the word dependencies of the sentence, inspired by the success achieved by [10,11], we first construct the graph of convolutional networks for each input sentence over the dependency tree.[2] After that, an adjacency matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ of the graph of the sentence is derived as follow:

$$D_{i,j} = \begin{cases} 1 & \text{if } w_i, w_j \text{ contains dependency} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

[2] In this work, we use spaCy toolkit to derive dependency tree of the sentence: https://spacy.io/.

**Table 1**
Examples of affective words from SenticNet.

| Word | SenticNet(word) |
|---|---|
| Great | 0.857 |
| Fantastic | 0.870 |
| Food | 0.170 |
| Bad | −0.800 |
| Balefully | −0.810 |
| Hard | −0.059 |

Here, following the previous GCN-based model [11], we construct the graph with un-directional, since we argue that the dependencies between words are mutual. We show a sketch of constructing directional and un-directional dependency graph in Fig. 2. In which, the un-directional dependency graph considers that parent nodes are equally influenced by their children nodes, that is, $D_{i,j} = D_{j,i}$.

### 3.5. Enhancing graph based on SenticNet

To leverage the affective information between contextual words and aspect words, we further enhance the representation of the adjacency matrix by incorporating the affective score from SenticNet [51]:

$$S_{i,j} = SenticNet(w_i) + SenticNet(w_j) \tag{2}$$

where $SenticNet(w_i) \in [-1, 1]$ represents the affective score of word $w_i$ in SenticNet. In which, $SenticNet(w_i) = 0$ denotes the word $w_i$ is a neutral word or inexistent in SenticNet. Then, we can extract the affective dependency between parent word and the child word in the dependency tree of the sentence. That is, the sentiment information between two dependent words is derived by the sum of their affective scores, i.e., the model tends to learn sentiment information from the words with the more intense sentiment. Here, we extract 39,891 words and their affective scores from SenticNet 6 in this work. The examples of words and their corresponding affective score are shown in Table 1.

In addition, existing GCN-based models for aspect sentiment analysis generally ignore to pay significant attention to the given aspect when constructing the graph. Hence, in this work, we further enhance the affective dependencies between contextual words and aspect words based on SenticNet:

$$T_{i,j} = \begin{cases} 1 & \text{if } w_i \text{ or } w_j \text{ is a aspect word} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

After that, we can obtain the enhanced adjacency matrix of the sentence:

$$A_{i,j} = D_{i,j} \times (S_{i,j} + T_{i,j} + 1) \tag{4}$$

The procedure of generating the adjacency matrix for each sentence is depicted in Algorithm 1.

### 3.6. SenticNet-based graph convolutional network

Based on the affective adjacency matrix derived over dependency tree and SenticNet, we feed the graph into the GCN layers to learn the affective dependencies for the given aspect. Then each node in the $l$th GCN layer is updated according to the hidden representations of its neighborhoods:

$$\mathbf{h}_i^l = relu(\tilde{\mathbf{A}}_i \mathbf{g}_i^{l-1} \mathbf{W}^l + \mathbf{b}^l) \tag{5}$$

$$\mathbf{g}_i^{l-1} = \mathcal{F}(\mathbf{h}_i^{l-1}) \tag{6}$$

**Algorithm 1** The procedure of deriving the adjacency matrix for each sentence

**Require:** a sequence of words $s = \{w_1, w_2, \cdots, w_n\}$; a set of aspect words $a = \{w_{a1}, w_{a2}, \cdots, w_{ak}\}$; a set of affective words generated from SenticNet; the dependency tree of the sentence *dependency*($s$).

**for** $i = 1 \rightarrow n$ **do**
  **for** $j = 1 \rightarrow n$ **do**
    **if** $dependency(w_i, w_j) \in dependency(s)$ **or** $i = j$ **then**
      ▷ Generated by dependency tree
      $D_{i,j} \leftarrow 1$
      ▷ Enhanced by SenticNet
      $S_{i,j} \leftarrow SenticNet(w_i) + SenticNet(w_j)$
      **if** $w_i \in a$ **or** $w_j \in a$ **then**
        $T_{i,j} \leftarrow 1$
      **else**
        $T_{i,j} \leftarrow 0$
      **end if**
      $A_{i,j} \leftarrow D_{i,j} \times (S_{i,j} + T_{i,j} + 1)$
    **else**
      $A_{i,j} \leftarrow 0$
    **end if**
  **end for**
**end for**

where $\mathbf{g}^{l-1}$ is the hidden representation evolved from the preceding GCN layer. $\mathcal{F}(\cdot)$ is a position-aware transformation function, which is exploited in a previous GCN-based ABSA work [11]. $\tilde{\mathbf{A}}$ is a normalized symmetric adjacency matrix:

$$\tilde{\mathbf{A}}_i = \mathbf{A}_i / (\mathbf{E}_i + 1) \tag{7}$$

where $\mathbf{E}_i = \sum_{j=1}^{n} A_{i,j}$ is the degree of $\mathbf{A}_i$. Here, the original nodes are built based on the hidden representations learn by Bi-LSTM layers, which take the word embeddings of context as input:

$$\mathbf{H}^c = \{\mathbf{h}_1^c, \mathbf{h}_2^c, \ldots, \mathbf{h}_n^c\} = Bi\text{-}LSTM(\mathbf{x}) \tag{8}$$

### 3.7. Aspect-specific sentiment representation

To highlight the significant features of aspect words, we exploit aspect-specific masking to mask the non-aspect words of the output vectors learned by the final GCN layer and keep the aspect word representations unchanged:

$$\tilde{\mathbf{h}}_t = \begin{cases} \tilde{\mathbf{h}}_t & \text{if } \tau \le t < \tau + k \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{9}$$

where $\tilde{\mathbf{h}}_t$ is the representation of the $t$th word learned by the final output of GCN layers, $\tau$ is the beginning index of aspect words in the sentence, and $k$ is the length of aspect words. We thus obtain the final representation of the aspect-specific masking:

$$\tilde{\mathbf{H}}_{mask} = \{\mathbf{0}, \ldots, \tilde{\mathbf{h}}_\tau, \ldots, \tilde{\mathbf{h}}_{\tau+k-1}, \ldots, \mathbf{0}\} \tag{10}$$

Based on the final representations learned via GCN layers, inspired by [11], we adopt a retrieval-based attention mechanism to derive significant features from the contextual affective words according to the specific aspect:

$$\beta_t = \sum_{i=1}^{n} \mathbf{h}_t^{c\top} \tilde{\mathbf{h}}_i = \sum_{i=\tau}^{\tau+k-1} \mathbf{h}_t^{c\top} \tilde{\mathbf{h}}_i \tag{11}$$

$$\alpha_t = \frac{exp(\beta_t)}{\sum_{i=1}^{n} exp(\beta_i)} \tag{12}$$

Here, the retrieval-based attention mechanism is adopted to measure the affective information and semantic relatedness between

**Table 2**
Statistics of the datasets.

| Dataset | Positive | | Neural | | Negative | |
|---|---|---|---|---|---|---|
| | Train | Test | Train | Test | Train | Test |
| Rest14 | 2164 | 728 | 637 | 196 | 807 | 196 |
| Lap14 | 994 | 341 | 464 | 169 | 870 | 128 |
| Rest15 | 1178 | 439 | 50 | 35 | 382 | 328 |
| Rest16 | 1620 | 597 | 88 | 38 | 709 | 190 |

contextual words and aspect words, lies in that constructing graph based on affective commonsense knowledge (i.e., SenticNet) can derive affective dependencies of the sentence, and further aspect-specific masking operation can obviously highlight the significance of aspect words. Hence, the final representation of the input vectors can be computed as follow:

$$\mathbf{r} = \sum_{i=1}^{n} \alpha_i \mathbf{h}_i^c \tag{13}$$

$$\mathbf{y} = softmax(\mathbf{W}_o \mathbf{r} + \mathbf{b}_o) \tag{14}$$

where $softmax(\cdot)$ is the softmax function to learn the output distribution of the sentiment classifier.

### 3.8. Model training

We adopt standard gradient descent algorithm to optimize and update the parameters of the proposed model. The objective to train the model is defined as minimizing the cross-entropy loss with $L_2$ regularization:

$$\mathcal{L} = -\sum_{i=1}^{S}\sum_{j=1}^{C} \hat{y}_i^j log y_i^j + \lambda \|\Theta\|^2 \tag{15}$$

where $S$ is the number of training samples, $C$ is the number of classes. $\hat{y}$ is the correct distribution of sentiment. $\Theta$ denotes all trainable parameters. $\lambda$ is the coefficient of $L_2$ regularization term.

## 4. Experiments

### 4.1. Datasets and experimental settings

We conduct experiments on four public benchmark datasets from restaurants and laptops domain of SemEval 2014 task 4 [54] (Rest14, Lap14), restaurants domain of SemEval 2015 task 12 [2] (Rest15) and restaurants domain of SemEval 2016 task 5 [55] (Rest16). Each sample consists of the review sentence, the aspect which consisting of one or multiple words, and the sentiment polarity towards the aspect.[3] The statistics of the datasets are shown in Table 2.

In our experiments, for all non-BERT based models, we use GloVe vectors [52] to map each word into 300-dimensional embedding. We set the depth of GCN layers as 2 in the model architecture. The coefficient $\lambda$ of $L_2$ regularization item is 0.00001. The dimensionality of hidden state vectors is set to 300 for all neural network layers. Adam with 0.001 learning rate is adopted to optimize and update the parameters of the model. We randomly initialize $\mathbf{W}$ and $\mathbf{b}$ with uniform distribution for all neural network layers. For the BERT-based models, we adopts "[CLS] sentence [SEP] aspect [SEP]" as input since the contextual sentiment representations are important for detecting the sentiment of a particular aspect. The dimension of word embedding is set to 768 based on the pre-trained uncased BERT-base model [53]. The learning rate is set to 0.00002. The batch size is set to 16 for non-BERT and BERT-based models. Inspired by [11], we report the experimental results of averaging 10 runs with random initialization.

[3] We removed sentence without explicit aspect or aspects in the same sentence contain different sentiment polarities.

### 4.2. Comparison models

To purely evaluate the performance of SenticNet for the task of ABSA based on GCN, we present another model that merely exploits affective words extracted from SenticNet to formulate the adjacency matrix of the graph without neither utilizing dependency tree nor enhancing the dependencies via the specific aspect, called Affective GCN. We compared the proposed models (Sentic GCN and Affective GCN) with the following methods:

- **TD-LSTM** [13] uses two LSTM networks to capture left and right directional contextual dependencies for a specific aspect.
- **ATAE-LSTM** [8] computes attention scores according to the specific aspect, and therefore model can pay attention to the key parts of the sentence.
- **MemNet** [56] exploits a multi-hop memory network to focus on aspect words in the sentence via word and position attention.
- **IAN** [9] interactively learns the relationships between contextual words and aspect words via interactive attention networks.
- **RAM** [14] explores a recurrent attention-based memory network for aspect sentiment prediction, which can capture long distance sentiment features for the aspects.
- **GCAE** [17] uses a gated CNN model to selectively control the propagation of the sentiment features with respect to the given aspect.
- **IARM** [4] extracts the influence of the preceding aspects via memory networks, which incorporates the neighboring aspects related information into the sentiment classification.
- **MGAN** [16] learns the interaction between aspect words and contextual words via exploring the fine-grained and coarse-grained attention mechanisms.
- **AOA** [57] learns the interaction between aspect words and contextual words, and focuses on the important parts in sentences via an attention-over-attention model.
- **TNet-LF** [58] adopts CNN to handle target-level sentiment classification and preserve and strengthen the informative part of contexts for the specific aspect.
- **TransCap** [23] transfers the sentiment features learned from document-level knowledge to aspect-level representations by adopting a transfer capsule network model.
- **IACapsNet** [22] adopts a capsule network to model vector-based feature representation and cluster the sentiment features by an EM routing algorithm for a specific aspect.
- **TD-GAT** [29] exploits a target-dependent graph attention network to extract the dependency relationships among words.
- **ASGCN-DT** [11] constructs a directional graph for each sentence over the dependency tree and then extracts syntactical information and word dependencies by a GCN.
- **ASGCN-DG** [11] is identical to ASGCN-DT, except the adjacency matrix of the graph is un-directional.
- **CDT** [10] extracts the dependencies between contextual words and aspect words by utilizing a GCN model over the dependency tree of the sentence.
- **DGEDT** [30] learns the flat representation and graph-based representation by a dependency graph enhanced dual-transformer network.

- **R-GAT** [31] explores a relational graph attention network to encode the aspect-based tree structure for sentiment prediction.
- **BERT** [53] is the vanilla BERT model, which adopts "[CLS] sentence [SEP] aspect [SEP]" as input.
- **SA-GCN+BERT** [59] is a GCN model based on BERT and the dependency tree, which employs the selective attention find important words to derive the representations of aspects.
- **TD-BERT** [60] is a BERT-based model for target-dependent sentiment analysis.
- **DGEDT-BERT** [30] is identical to DGEDT, but adopted BERT as the aspect-based encoder.
- **R-GAT+BERT** [31] is identical to R-GAT, but adopted BERT as the aspect-based encoder.
- **LCFS-ASC** [7] explores the grammatical aspect of the sentence and employs the self-attention mechanism for syntactical learning based on BERT for ABSA.
- **SGGCN+BERT** [61] regulates the hidden vectors of the graph-based models for using the information from the aspects, and computes a gate vector for each layer to leverage the representations of the aspects.
- **Affective GCN** is our proposed model, which merely uses the affective words extracted from SenticNet to construct an un-directional graph for each sentence.
- **Sentic GCN** is our complete proposed model, which construct the graph for each sentence over the dependency tree, and then enhance the dependencies between contextual words and aspect words via leveraging the affective information of SenticNet.
- **Sentic GCN-D** is identical to our proposed Sentic GCN model, but utilizes directional graphs in the GCN, whose adjacency matrices are more sparse than the un-directional graphs.
- **Sentic GCN-BERT** is identical to our proposed Sentic GCN model, but replaces the embedding module and the LSTM-layers with the uncased basic pre-trained BERT. That is the GCN-layers in Sentic GCN-BERT takes the output encodings of the sentence from BERT and the corresponding graph as input for learning graph representations.

### *4.3. Main results*

As shown in Table 3, the experimental results demonstrate that our proposed Sentic GCN performs better than all comparison models on four benchmark datasets, including deep neural networks, graph networks and BERT-based models. This verifies the effectiveness of the proposed model in the task of ABSA. Specifically, the proposed Sentic GCN performs significantly better than the previous GCN-based models that construct graphs over the dependency tree (TD-GAT, ASGCN-DT, ASGCN-DG, CDT, DGEDT and R-GAT), verifying the effectiveness and desirability of enhancing word dependencies of the sentence via leveraging the affective information from SenticNet. In which, compared with the previous remarkable GCN-based model (R-GAT), our proposed Sentic GCN can also achieves comparative performance in relatively poor result (i.e. Macro-F1 score on Rest14 dataset). In addition, compared with Sentic GCN-D whose graphs of sentences are directional, the proposed Sentic GCN achieves better performance on all the four datasets. A possible reason is that the dependencies between parents nodes and children nodes are mutual and equally significant, hence constructing an un-directional graph for convolutional networks can take full advantage of the dependency tree of the sentence and derive more precise affective relations in comparison with the directional one. It is also noteworthy that the proposed Affective GCN, which merely produces the graph of the sentence based on the affective words from SenticNet, performs better than ASGCN-DT and ASGCN-DG and achieves comparable results with the other more powerful graph-based models. This implies that affective information retrieved SenticNet can improve the learning ability of sentiment extraction of GCN model in ABSA.

On the other hand, for BERT-based models, our proposed Sentic GCN-BERT significantly outperforms both the vanilla BERT model and several remarkable BERT-based models on Rest15 and Rest16 datasets, and performs competitively with the previous remarkable SGGCN+BERT and overall better than the other BERT-based baseline models on Rest14 and Lap14 datasets. The results demonstrate the effectiveness of our proposed graph construction based on SenticNet for GCN model when employing more powerful pre-trained encoder in the task of ABSA.

### *4.4. Ablation study*

To investigate the impact of different components of constructing graphs for the proposed Sentic GCN brings to the performance of ABSA, we conduct an experiment of ablation study and demonstrate the results in Table 4. We can observe from Table 4, the model that utilizes neither "SenticNet" nor "aspect" to construct the graph for the sentence (i.e., Sentic GCN w/o SenticNet+aspect) performs unsatisfactorily on all datasets. Comparatively, the model without "dependency" and "aspect" (i.e., Sentic GCN w/o dependency+aspect) evidently performs better. This indicates SenticNet can provide affective information into the dependencies between context and aspects, and therefore more precise sentiment features are derived to improve the performance of sentiment prediction for specific aspects. In addition, both of removal of "SenticNet" (Sentic GCN w/o SenticNet) and removal of "dependency" (Sentic GCN w/o dependency) lead to performance drops. We can conclude that both "SenticNet" and "dependency" are important to produce graphs for convolutional computation, in particular, exploiting SenticNet to enhance the dependency tree of a sentence can promote the model to leverage affective relations for predicting sentiment with respect to the specific aspect to a great extent. We can also notice that removal of "aspect" reduces the performance of the model slightly, which broadly indicates that incorporating aspect information into the graph can stimulate the model to focus on the important parts of the sentence in a more explicable way, so as to achieve the superior performance approaching the complete Sentic GCN model.

### *4.5. Impact of SenticNet*

To further demonstrate that SenticNet can enhance the word dependency representations of the sentence and improve the performance of predicting aspect-specific sentiment, we conduct experiments by employing different proportions of affective words from SenticNet on Rest15 dataset with Sentic GCN. The results are shown in Fig. 3. We can observe that no matter in any proportion of SenticNet, the performance of Sentic GCN is better than abandoning SenticNet (blue dotted lines) on both Accuracy and Macro-F1, which further verifies the effectiveness of SenticNet in predicting aspect-based sentiment polarity with GCN. It is also noteworthy that the performance of model fluctuates on both Accuracy and Macro-F1 scores when the proportion of SenticNet is less than 60%, while the performance of Sentic GCN presents steady upward trends with the increasing number of affective words from SenticNet when the proportion of SenticNet is large ($\geq$60%). This implies that the more sufficient affective words can better enhance the word dependencies of sentences and derive more precise dependency graphs. In addition, Rest15 performs better based on sorted SenticNet in comparison with random selection when the proportion of SenticNet is more than 60%, which denotes the higher score affective words can better exhibit the sentiment information of the sentence.

**Table 3**
Main experimental results on four datasets. Acc. represents accuracy, F1 represents Macro-F1 score. Best results are in bold face and second best underlined. The results with ♮ are retrieved from [11], with † are retrieved from the original papers, and others are reported based on the open source codes.

| Model | | Rest14 (%) | | Lap14 (%) | | Rest15 (%) | | Rest16 (%) | |
|---|---|---|---|---|---|---|---|---|---|
| | | Acc. | F1 | Acc. | F1 | Acc. | F1 | Acc. | F1 |
| Network | TD-LSTM [13] | 78.00† | 66.73† | 71.83† | 68.43† | 76.39 | 58.70 | 82.16 | 54.21 |
| | ATAE-LSTM [8] | 77.20† | 67.02 | 68.70† | 63.93 | 78.48 | 60.53 | 83.77 | 61.71 |
| | MemNet [56]♮ | 79.61 | 69.64 | 70.64 | 65.17 | 77.31 | 58.28 | 85.44 | 65.99 |
| | IAN [9]♮ | 79.26 | 70.09 | 72.05 | 67.38 | 78.54 | 52.65 | 84.74 | 55.21 |
| | RAM [14] | 80.23† | 70.80† | 74.49† | 71.35† | 79.98 | 60.57 | 83.88 | 62.14 |
| | GCAE [17] | 77.28† | 62.45 | 69.14† | 68.71 | 77.56 | 56.03 | 83.70 | 62.69 |
| | IARM [4]† | 80.00 | – | 73.80 | – | – | – | – | – |
| | MGAN [16] | 81.25† | 71.94† | 75.39† | 72.47† | 79.36 | 57.26 | 87.06 | 62.29 |
| | AOA [57]♮ | 79.97 | 70.42 | 72.62 | 67.52 | 78.17 | 57.02 | 87.50 | 66.21 |
| | TNet-LF [58]♮ | 80.42 | 71.03 | 74.61 | 70.14 | 78.47 | 59.47 | 89.07 | 70.43 |
| | TransCap [23]† | 79.29 | 70.85 | 73.87 | 70.10 | – | – | – | – |
| | IACapsNet [22]† | 81.79 | 73.40 | 76.80 | 73.29 | – | – | – | – |
| Graph | TD-GAT [29]† | 81.10 | – | 73.70 | – | – | – | – | – |
| | ASGCN-DT [11]♮ | 80.86 | 72.19 | 74.14 | 69.24 | 79.34 | 60.78 | 88.69 | 66.64 |
| | ASGCN-DG [11]♮ | 80.77 | 72.02 | 75.55 | 71.05 | 79.89 | 61.89 | 88.99 | 67.48 |
| | CDT [10]† | 82.30 | 74.02 | 77.19 | 72.99 | – | – | 85.58 | 69.93 |
| | DGEDT [30]† | <u>83.90</u> | 75.10 | 76.80 | 72.30 | <u>82.10</u> | 65.90 | <u>90.80</u> | <u>73.80</u> |
| | R-GAT [31] | 83.30† | **76.08**† | <u>77.42</u>† | <u>73.76</u>† | 80.83 | 64.17 | 88.92 | 70.89 |
| Ours | Affective GCN | 81.62 | 73.86 | 77.30 | 73.15 | 81.23 | <u>66.28</u> | 89.72 | 73.30 |
| | Sentic GCN-D | 81.78 | 73.96 | 76.02 | 72.08 | 80.78 | 64.75 | 89.56 | 71.94 |
| | Sentic GCN | **84.03** | <u>75.38</u> | **77.90** | **74.71** | **82.84** | **67.32** | **90.88** | **75.91** |
| Bert | BERT [53] | 84.11 | 76.68 | 77.59 | 73.28 | 83.48 | 66.18 | 90.10 | 74.16 |
| | SA-GCN+BERT [59]† | 85.80 | 79.70 | 81.70 | 78.80 | – | – | – | – |
| | TD-BERT [60]† | 85.10 | 78.35 | 78.87 | 74.38 | – | – | – | – |
| | DGEDT-BERT [30]† | 86.30 | 80.00 | 79.80 | 75.60 | <u>84.00</u> | <u>71.00</u> | <u>91.90</u> | <u>79.00</u> |
| | R-GAT+BERT [31] | 86.60† | <u>81.35</u>† | 78.21† | 74.07† | 83.22 | 69.73 | 89.71 | 76.62 |
| | LCFS-ASC [7]† | 86.71 | 80.31 | 80.52 | 77.13 | – | – | – | – |
| | SGGCN+BERT [61] | **87.20**† | **82.50**† | **82.80**† | **80.20**† | 82.72 | 65.86 | 90.52 | 74.53 |
| Ours | Sentic GCN-BERT | <u>86.92</u> | 81.03 | <u>82.12</u> | <u>79.05</u> | **85.32** | **71.28** | **91.97** | **79.56** |

**Table 4**
Experimental results of ablation study on four datasets. "s" represents SenticNet, "d" represents dependency tree, "a" represents highlighting the word dependencies according to the specific aspect.

| Model | Rest14 (%) | | Lap14 (%) | | Rest15 (%) | | Rest16 (%) | |
|---|---|---|---|---|---|---|---|---|
| | Acc. | F1 | Acc. | F1 | Acc. | F1 | Acc. | F1 |
| Sentic GCN w/o s+a | 80.56 | 71.86 | 75.21 | 70.82 | 79.35 | 61.80 | 88.12 | 69.20 |
| Sentic GCN w/o d+a | 81.42 | 72.93 | 76.80 | 72.72 | 80.72 | 64.73 | 88.47 | 71.08 |
| Sentic GCN w/o s | 80.92 | 72.13 | 75.68 | 71.24 | 80.15 | 62.28 | 88.75 | 70.32 |
| Sentic GCN w/o d | 82.51 | 73.88 | 76.02 | 71.67 | 80.73 | 64.57 | 88.93 | 71.28 |
| Sentic GCN w/o a | 82.92 | 74.12 | 76.35 | 72.73 | 81.03 | 65.36 | 89.52 | 72.75 |
| Sentic GCN | **84.03** | **75.38** | **77.90** | **74.71** | **82.84** | **67.32** | **90.88** | **75.91** |

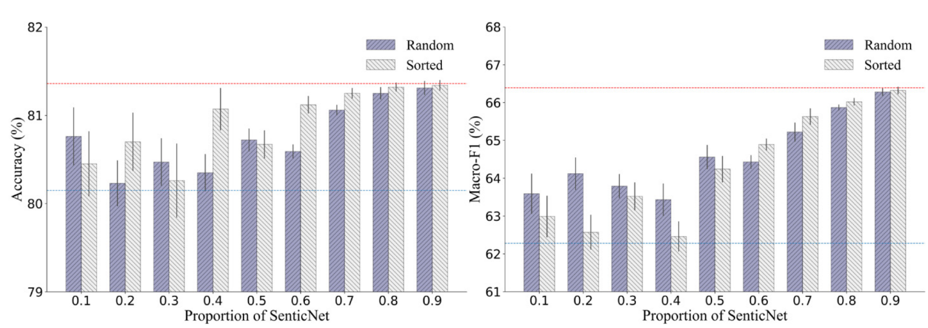



**Fig. 3.** The impact of the proportion of SenticNet. Accuracy and Macro-F1 scores based on different proportions of SenticNet are reported. Blue dotted lines represent the result of using 0% SenticNet. Red dotted lines represent the result of using 100% SenticNet. "Random" denotes we select affective words from SenticNet randomly. "Sorted" represents we preferentially select affective words whose absolute values of affective scores are higher.

### 4.6. Impact of GCN layers

We analyze the impact of the depth of the proposed Sentic GCN in this section. We vary the number of GCN layers from 1 to 8 and show the experimental results in Fig. 4. Note that 2 layers GCN achieves overall better performance than other numbers of GCN layers, and thus the number of GCN layers is set to 2 in our experiments. The one-layer Sentic GCN performs unsatisfactorily on all the four datasets for both accuracy and Macro-F1 scores, which implies that one-layer GCN is insufficient to leverage the affective dependencies of the sentence with respect to the specific aspect. In addition, the performance of Sentic GCN fluctuates with

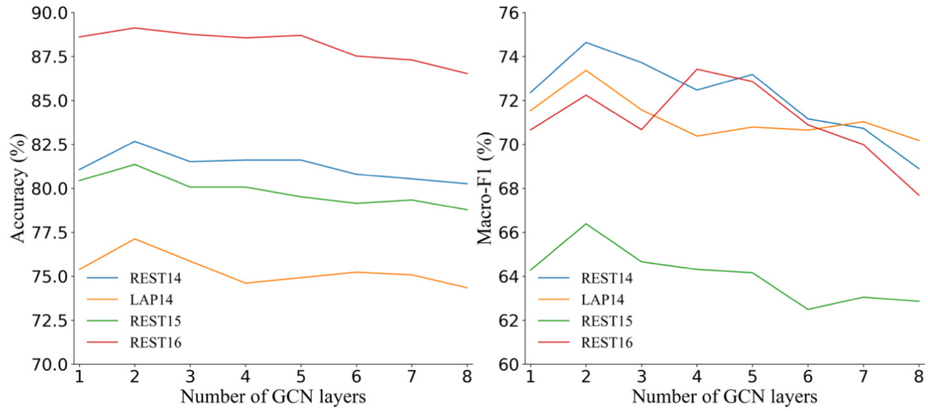



**Fig. 4.** The impact of the number of GCN layers of the proposed Sentic GCN. Accuracy and Macro-F1 scores based on different numbers of GCN layers are reported.

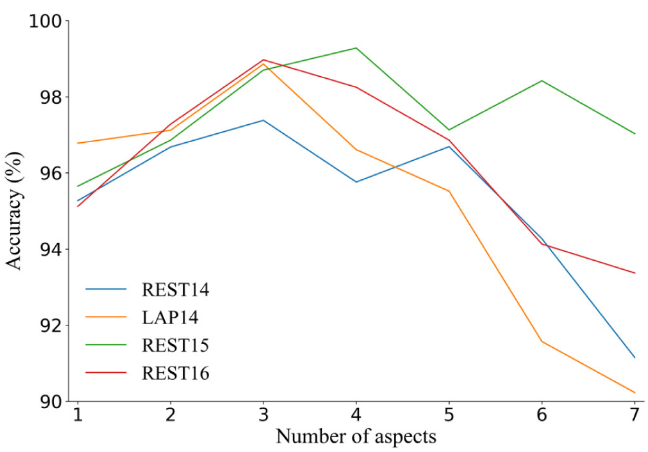



**Fig. 5.** The impact of the number of aspects in the sentences.

the increasing number of GCN layers and essentially tends to decline when the model depth is greater than 2. This shows that roughly increasing the number of GCN layers may attenuate the learning of the model due to the sharp increase of parameters.

### 4.7. Impact of multiple aspects

As mentioned above, the key challenge of ABSA is to distinguish different sentiment polarities for different aspects in the same sentence. Hence, in this section, we further analyze the effects of the performance of the proposed Sentic GCN brought by multiple aspects sentences that contain different numbers of aspects. We separate the training instances into different groups according to the number of aspects in the sentences and report the training accuracy for different data groups.[4] The results are demonstrated in Fig. 5. We can observe from Fig. 5, when the number of aspects in the sentences is less than 3, the proposed Sentic GCN shows tendencies to ascend on all datasets with the increase of the number. One main reason may be that sentences contain few aspects relatively produce simple dependency graphs, and thus our proposed Sentic GCN is more expert in extracting the contextual affective dependencies according to the aspects. By contrast, increasing the number of aspects from 3 to 7 leads to performance fluctuated drops, but still, achieve reasonable performance on all datasets (on REST15 in particular). This implies that multiple-aspect sentences partly impede the learning of the model in comparison with those straightforward instances, however, thanks to the advantage of affective dependency graphs, the proposed Sentic GCN can still cope with the intricate dependencies of the multiple-aspect sentences (the accuracy is over 90% on all datasets).

[4] Here, we ignore the sentences whose aspects numbers are more than 7 since the count of these instances is too small for any meaningful comparison.

### 4.8. Visualizations

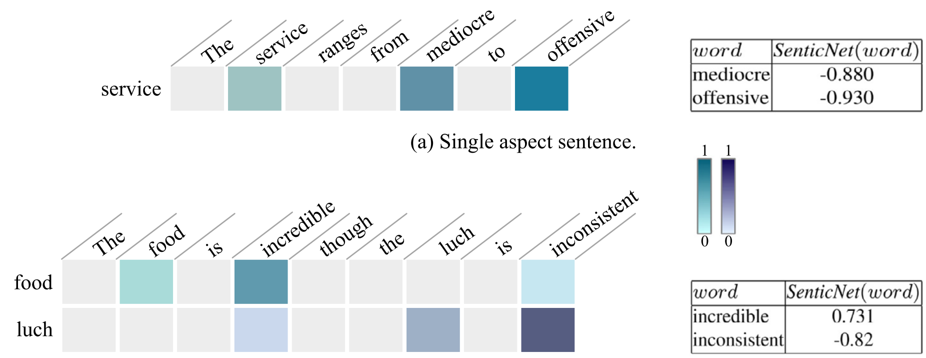
To qualitatively demonstrate how the proposed Sentic GCN model improves the performance of predicting aspect-based sentiment polarity, we visualize the attention weights by showing two typical examples (i.e., a single-aspect instance and a multiple-aspects instance). The results and corresponding affective words are demonstrated in Fig. 6. We can see that the model can pay more attention to the key affective words for extracting sentiment features corresponding to specific aspects. In particular, the multiple-aspects example denotes that the proposed Sentic GCN can distinguish different aspects and focus on different contextual affective words according to different aspects. This indicates that enhancing word dependencies based on SenticNet and the specific aspect can derive precise aspect-specific sentiment features and improve the performance of ABSA.

### 4.9. Case study

To better analyze how Sentic GCN works in extracting key affective words for the specific aspect with the help of SenticNet, we present a case study by reporting the learning of some typical testing examples. The results are demonstrated in Fig. 7. The sentiment expressed towards the aspect is vague in Review1, hence ASGCN reveals a wrong prediction. While owing to the enhancement of SenticNet, our proposed Sentic GCN can notice both the key positive and negative words for predicting a *neutral* sentiment polarity of the given aspect. For multiple aspect sentences, Review2 is a sentence that contains diverse affective words, i.e., "small", "lovely" and "helpful". In spite of learning opposed affective words, Sentic GCN can still extract the most important sentiment information for distinct aspects. In Review3, sentiment information of aspect "mioposto" is obscure. ASGCN cannot obtain a correct result for aspect "mioposto", while Sentic GCN output an accurate sentiment polarity for the aspect profit from the negative affective word "only".

## 5. Conclusion

To develop the merit of affective commonsense knowledge in the task of aspect-based sentiment analysis, we propose an affective knowledge-enhanced graph model, called Sentic GCN, which




(a) Single aspect sentence.

(b) Multiple aspects sentence.

**Fig. 6.** The attention visualizations of typical samples.

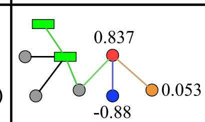
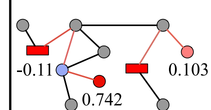
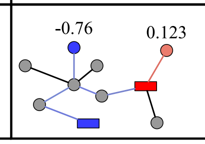
| Reviews | ASGCN-DG | Sentic GCN | SenticNet | Sentic Graph |
|---|---|---|---|---|
| 1. The [regular menu] here is ***slightly*** ***above*** ***average*** | (regular menu, positive) | (regular menu, neutral) | (slightly, -0.88) (above, 0.837) (average, 0.053) |   |
| 2. The [space] is ***small*** and ***lovely***, and the [service] is ***helpful*** | (space, negative) (service, positive) | (space, positive) (service, positive) | (small, -0.11) (lovely, 0.742) (helpful, 0.103) |   |
| 3. The ***only*** positive thing about [mioposto] is the ***nice*** [location] | (mioposto, positive) (location, positive) | (mioposto, negative) (location, positive) | (only, -0.76) (nice, 0.123) |   |

**Fig. 7.** Case Study. Typical examples learned by our Sentic GCN and comparison model (ASGCN-DG) are reported. Red, blue, and green respectively represent positive, negative, and neutral sentiment. In the "Sentic Graph" column, circles represent contextual words, rectangles represent aspect words. Lines in darker colors represent stronger sentiment.

explores a novel solution to construct graphs for convolutional networks based on SenticNet, a widely used external affective commonsense knowledge base. To be specific, we first produce graphs over dependency trees of sentences. And further, we enhance the word dependencies of each sentence by incorporating affective information into the graph and highlight the specific aspect via leveraging the relations between aspect words and contextual affective words. Consequently, the proposed model can focus on the key affective words and derive precise sentiment features corresponding to different aspects. Experimental results on four benchmark datasets show that the proposed model can enhance the representations of contextual word dependencies and achieves state-of-the-art performance.

## CRediT authorship contribution statement

**Bin Liang:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Hang Su:** Software, Data curation, Investigation, Validation. **Lin Gui:** Methodology, Writing – original draft, Writing – review & editing. **Erik Cambria:** Conceptualization, Methodology, Resources, Writing – original draft. **Ruifeng Xu:** Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

## Declaration of competing interest

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

## Acknowledgments

This work was partially supported by National Natural Science Foundation of China (61632011, 61876053, 62006062), Guangdong Province Covid-19 Pandemic Control Research, China Funding (2020KZDZX1224), Shenzhen Foundational Research, China Funding (JCYJ20180507183527919, JCYJ20180507183608379), the Joint Lab of China Merchants Securities and HITSZ, and the Agency for Science, Technology and Research (A*STAR), Singapore under its AME Programmatic, Singapore Funding Scheme (Project #A18A2b0046).

## References

[1] B. Liu, L. Zhang, A survey of opinion mining and sentiment analysis, in: Mining Text Data, Springer, 2012, pp. 415–463.

[2] M. Pontiki, D. Galanis, H. Papageorgiou, S. Manandhar, I. Androutsopoulos, SemEval-2015 task 12: Aspect based sentiment analysis, in: Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), 2015, pp. 486–495.

[3] J. Du, L. Gui, Y. He, R. Xu, X. Wang, Convolution-based neural attention with applications to sentiment classification, IEEE Access 7 (2019) 27983–27992.

[4] N. Majumder, S. Poria, A. Gelbukh, M.S. Akhtar, E. Cambria, A. Ekbal, IARM: Inter-aspect relation modeling with memory networks in aspect-based sentiment analysis, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 3402–3411.

[5] Y. Ma, H. Peng, T. Khan, E. Cambria, A. Hussain, Sentic LSTM: a hybrid network for targeted aspect-based sentiment analysis, Cogn. Comput. 10 (4) (2018) 639–650.

[6] J. Tang, Z. Lu, J. Su, Y. Ge, L. Song, L. Sun, J. Luo, Progressive self-supervised attention learning for aspect-level sentiment analysis, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 557–566.
[7] M.H. Phan, P.O. Ogunbona, Modelling context and syntactical features for aspect-based sentiment analysis, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 3211–3220.
[8] Y. Wang, M. Huang, x. zhu, L. Zhao, Attention-based LSTM for aspect-level sentiment classification, in: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, pp. 606–615.
[9] D. Ma, S. Li, X. Zhang, H. Wang, Interactive attention networks for aspect-level sentiment classification, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, 2017, pp. 4068–4074.
[10] K. Sun, R. Zhang, S. Mensah, Y. Mao, X. Liu, 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), 2019, pp. 5679–5688.
[11] C. Zhang, Q. Li, D. Song, Aspect-based sentiment classification with aspect-specific graph convolutional 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), 2019, pp. 4567–4577.
[12] E. Cambria, A. Hussain, C. Havasi, C. Eckl, Common sense computing: From the society of mind to digital intuition and beyond, in: J. Fierrez, J. Ortega, A. Esposito, A. Drygajlo, M. Faundez-Zanuy (Eds.), Biometric ID Management and Multimodal Communication, in: Lecture Notes in Computer Science, vol. 5707, Springer, Berlin Heidelberg, 2009, pp. 252–259.
[13] D. Tang, B. Qin, X. Feng, T. Liu, Effective LSTMs for target-dependent sentiment classification, in: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, 2016, pp. 3298–3307.
[14] P. Chen, Z. Sun, L. Bing, W. Yang, Recurrent attention network on memory for aspect sentiment analysis, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017, pp. 452–461.
[15] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, Exploiting document knowledge for aspect-level sentiment classification, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2018, pp. 579–585.
[16] F. Fan, Y. Feng, D. Zhao, Multi-grained attention network for aspect-level sentiment classification, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 3433–3442.
[17] W. Xue, T. Li, Aspect based sentiment analysis with gated convolutional networks, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, 2018, pp. 2514–2523.
[18] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, An interactive multi-task learning network for end-to-end aspect-based sentiment analysis, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 504–515.
[19] D. Hazarika, S. Poria, P. Vij, G. Krishnamurthy, E. Cambria, R. Zimmermann, Modeling inter-aspect dependencies for aspect-based sentiment analysis, in: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), 2018, pp. 266–270.
[20] Z. Chen, T. Qian, Relation-aware collaborative learning for unified aspect-based sentiment analysis, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 3685–3694.
[21] W. Zhao, H. Peng, S. Eger, E. Cambria, M. Yang, Towards scalable and reliable capsule networks for challenging NLP applications, in: ACL, 2019, pp. 1549–1559.
[22] C. Du, H. Sun, J. Wang, Q. Qi, J. Liao, T. Xu, M. Liu, Capsule network with interactive attention for aspect-level sentiment classification, 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), 2019, pp. 5488–5497.
[23] Z. Chen, T. Qian, Transfer capsule network for aspect level sentiment classification, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 547–556.
[24] L. Yao, C. Mao, Y. Luo, Graph convolutional networks for text classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 7370–7377.
[25] L. Huang, D. Ma, S. Li, X. Zhang, H. Wang, Text level graph neural network for text classification, 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), 2019, pp. 3444–3450.
[26] Y. Zhang, P. Qi, C.D. Manning, Graph convolution over pruned dependency trees improves relation extraction, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 2205–2215.
[27] R. Ding, P. Xie, X. Zhang, W. Lu, L. Li, L. Si, A neural multi-digraph model for Chinese NER with gazetteers, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 1462–1467.
[28] T. Gui, Y. Zou, Q. Zhang, M. Peng, J. Fu, Z. Wei, X. Huang, A lexicon-based graph neural network for Chinese NER, 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), 2019, pp. 1040–1050.
[29] B. Huang, K. Carley, 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), 2019, pp. 5469–5477.
[30] H. Tang, D. Ji, C. Li, Q. Zhou, Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 6578–6588.
[31] K. Wang, W. Shen, Y. Yang, X. Quan, R. Wang, Relational graph attention network for aspect-based sentiment analysis, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 3229–3238.
[32] T. Young, E. Cambria, I. Chaturvedi, H. Zhou, S. Biswas, M. Huang, Augmenting end-to-end dialogue systems with commonsense knowledge, in: AAAI, 2018, pp. 4970–4977.
[33] P. Yang, L. Li, F. Luo, T. Liu, X. Sun, Enhancing topic-to-essay generation with external commonsense knowledge, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 2019, pp. 2002–2012.
[34] P. Parthasarathi, J. Pineau, Extending neural generative conversational model using external knowledge sources, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, 2018, pp. 690–695.
[35] N. Ofek, S. Poria, L. Rokach, E. Cambria, A. Hussain, A. Shabtai, Unsupervised commonsense knowledge enrichment for domain-specific sentiment analysis, Cogn. Comput. 8 (3) (2016) 467–477.
[36] T. Mihaylov, A. Frank, Knowledgeable reader: Enhancing cloze-style reading comprehension with external commonsense knowledge, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 821–832.
[37] I. Chaturvedi, R. Satapathy, S. Cavallari, E. Cambria, Fuzzy commonsense reasoning for multimodal sentiment analysis, Pattern Recognit. Lett. 125 (264–270) (2019).
[38] E. Cambria, A. Hussain, C. Havasi, C. Eckl, Sentic computing: Exploitation of common sense for the development of emotion-sensitive systems, in: A. Esposito, N. Campbell, C. Vogel, A. Hussain, A. Nijholt (Eds.), Development of Multimodal Interfaces: Active Listening and Synchrony, in: Lecture Notes in Computer Science, Springer, Berlin, 2010, pp. 148–156.
[39] S. Poria, I. Chaturvedi, E. Cambria, F. Bisio, Sentic LDA: Improving on LDA with semantic similarity for aspect-based sentiment analysis, in: IJCNN, 2016, pp. 4465–4473.
[40] M. Dragoni, S. Poria, E. Cambria, OntoSenticNet: A commonsense ontology for sentiment analysis, IEEE Intell. Syst. 33 (3) (2018) 77–85.
[41] M. Dragoni, I. Donadello, E. Cambria, Ontosenticnet 2: enhancing reasoning within sentiment analysis, IEEE Intell. Syst. 36 (2021) 5.
[42] E. Cambria, T. Mazzocco, A. Hussain, C. Eckl, Sentic medoids: Organizing affective common sense knowledge in a multi-dimensional vector space, in: D. Liu, H. Zhang, M. Polycarpou, C. Alippi, H. He (Eds.), Advances in Neural Networks, in: Lecture Notes in Computer Science, vol. 6677, Springer-Verlag, 2011, pp. 601–610.
[43] E. Cambria, R. Speer, C. Havasi, A. Hussain, SenticNet: A publicly available semantic resource for opinion mining, in: AAAI CSK, Arlington, 2010, pp. 14–18.
[44] E. Cambria, C. Havasi, A. Hussain, SenticNet 2: A semantic and affective resource for opinion mining and sentiment analysis, in: FLAIRS, Marco Island, 2012, pp. 202–207.
[45] E. Cambria, D. Olsher, D. Rajagopal, SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis, in: AAAI, Quebec City, 2014, pp. 1515–1521.
[46] E. Cambria, S. Poria, R. Bajpai, B. Schuller, SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives, in: COLING, 2016, pp. 2666–2677.
[47] E. Cambria, S. Poria, D. Hazarika, K. Kwok, SenticNet 5: Discovering conceptual primitives for sentiment analysis by means of context embeddings, in: AAAI, 2018, pp. 1795–1802.

[48] E. Cambria, Y. Li, F.Z. Xing, S. Poria, K. Kwok, SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis, in: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, 2020, pp. 105–114.
[49] Y. Li, Q. Pan, T. Yang, S. Wang, J. Tang, E. Cambria, Learning word representations for sentiment analysis, Cogn. Comput. 9 (6) (2017) 843–851.
[50] F.Z. Xing, F. Pallucchini, E. Cambria, Cognitive-inspired domain adaptation of sentiment lexicons, Inf. Process. Manage. 55 (3) (2019) 554–564.
[51] Y. Ma, H. Peng, E. Cambria, Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM, in: Proceedings of the Thirty-Second Conference on Artificial Intelligence, (AAAI-18), 2018, pp. 5876–5883.
[52] J. Pennington, R. Socher, C. Manning, Glove: Global vectors for word representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532–1543.
[53] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 2019, pp. 4171–4186.
[54] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, SemEval-2014 task 4: Aspect based sentiment analysis, in: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEVal 2014), 2014, pp. 27–35.
[55] M. Pontiki, D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. AL-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, V. Hoste, M. Apidianaki, X. Tannier, N. Loukachevitch, E. Kotelnikov, N. Bel, S.M. Jiménez-Zafra, G. Eryiğit, SemEval-2016 task 5: Aspect based sentiment analysis, in: Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), 2016, pp. 19–30.
[56] D. Tang, B. Qin, T. Liu, Aspect level sentiment classification with deep memory network, in: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, pp. 214–224.
[57] B. Huang, Y. Ou, K.M. Carley, 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, 2018, pp. 197–206.
[58] X. Li, L. Bing, W. Lam, B. Shi, Transformation networks for target-oriented sentiment classification, in: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2018, pp. 946–956.
[59] X. Hou, J. Huang, G. Wang, K. Huang, X. He, B. Zhou, Selective attention based graph convolutional networks for aspect-level sentiment classification, 2019, arXiv preprint arXiv:1910.10857.
[60] Z. Gao, A. Feng, X. Song, X. Wu, Target-dependent sentiment classification with BERT, IEEE Access 7 (2019) 154290–154299.
[61] A.P.B. Veyseh, N. Nouri, F. Dernoncourt, Q.H. Tran, D. Dou, T.H. Nguyen, Improving aspect-based sentiment analysis with gated graph convolutional networks and syntax-based regulation, in: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings, 2020, pp. 4543–4548.