S^2GSL: Incorporating Segment to Syntactic Enhanced Graph Structure Learning for Aspect-based Sentiment Analysis

Bingfeng Chen[1], **Qihan Ouyang**[1], **Yongqi Luo**[1], **Boyan Xu**[1*], **Ruichu Cai**[1], **Zhifeng Hao**[1,2]
[1]School of Computer Science, Guangdong University of Technology
[2]College of Science, Shantou University
chenbingfeng@gdut.edu.cn
{ouyangqihan0720, lyongqi001, hpakyim, cairuichu}@gmail.com
haozhifeng@stu.edu.cn

Abstract

Previous graph-based approaches in Aspect-based Sentiment Analysis(ABSA) have demonstrated impressive performance by utilizing graph neural networks and attention mechanisms to learn structures of static dependency trees and dynamic latent trees. However, incorporating both semantic and syntactic information simultaneously within complex global structures can introduce irrelevant contexts and syntactic dependencies during the process of graph structure learning, potentially resulting in inaccurate predictions. In order to address the issues above, we propose S^2GSL, incorporating Segment to Syntactic enhanced Graph Structure Learning for ABSA. Specifically, S^2GSL is featured with a segment-aware semantic graph learning and a syntax-based latent graph learning enabling the removal of irrelevant contexts and dependencies, respectively. We further propose a self-adaptive aggregation network that facilitates the fusion of two graph learning branches, thereby achieving complementarity across diverse structures. Experimental results on four benchmarks demonstrate the effectiveness of our framework.

1 Introduction

Aspect-based Sentiment Analysis (ABSA) is a fine-grained sentiment analysis task that aims to recognize the sentiment polarities of multiple aspects within a given sentence. For example, in the sentence "The falafel was rather overcooked and dried but the chicken was fine," the sentiment polarity of the aspect words "falafel" and "chicken" is recognized as negative and positive, respectively. The ABSA task presents a notable challenge in accurately recognizing the sentiment polarity of specific aspect words, particularly when they are influenced by other aspect words with contrasting polarities within the given context.

* Corresponding authors

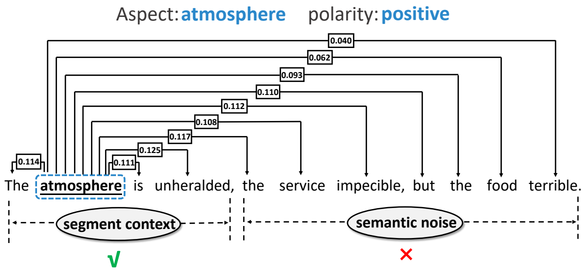


Figure 1: The attention weight of the aspect word "atmosphere" in relation to other words in the sentence."×" refers to noise information for "atmosphere".

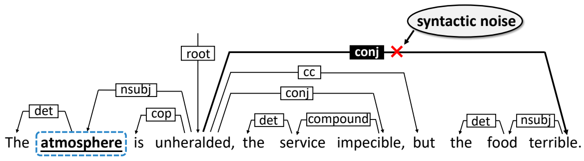


Figure 2: Dependency tree of an example sentence.

Leading graph-based approaches tackle ABSA tasks by learning either prior static structures and learnable dynamic semantic structures via graph neural networks and attention mechanisms. For instance, (Chen et al., 2020) proposed combining external syntactic dependency tree and implicit graph to generate aspect-specific representations using GCN; (Li et al., 2021) proposed to construct a SemGCN module and a SynGCN module using an attention mechanism and a syntactic dependency tree.

Although promising results were reported, we observe that existing graph structure learning approaches are still prone to incorrect predictions with hard samples containing multiple aspect words. Existing solutions introduce a heavy structure learning process leading to the following two main limitations: (i) Approaches reliant on attention mechanism are vulnerable to irrelevant context, potentially resulting in misalignment or weak linking. As shown in Figure 1, for the aspect word "atmosphere", except for the clause "The atmosphere is unheralded", the other parts are redundant

Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 13366–13379
August 11-16, 2024 ©2024 Association for Computational Linguistics

for judging sentiment polarity, i.e., semantic noise. Negatively influenced by irrelevant context, this results in only weak linking to "atmosphere" and the corresponding opinion word "unheralded". (ii) The global structure of the dependency tree for parsing complex long sentences cannot avoid containing irrelevant dependency information for polarity judgment. As shown in Figure 2, the "conj" relation connecting "unheralded" and "terrible" is the syntactic noise for the aspect “atmosphere". Thus, the key to tackling these limitations is efficiently dividing a complex sentence into multiple local clauses in the graph structure learning process.

In this paper, we propose S^2GSL, incorporating Segment to Syntactic enhanced Graph Structure Learning for ABSA. To minimize the negative impact of irrelevant structures, S^2GSL introduces constituent trees to decompose the complex structure of the input sentences. To illustrate the role of constituent trees, Figure 3 presents a constituent tree and its third-layer segment matrix. This segment matrix can divide a sentence into three semantically complete paragraphs, which help to facilitate the alignment between each aspect and its corresponding opinion and filter out the contextual information unrelated to the respective paragraph. Specifically, we devise a segment-aware semantic graph(SeSG) branch by using a supervised dynamic local attention on the constituent tree, to learn the local semantic structure of each aspect. Sharing the same idea with leading graph-based approaches, S^2GSL has also been designed with a syntax-based latent graph(SyLG) branch that utilizes syntactic dependency labels to enhance the latent tree construction. The difference from past work is that we introduce an attention-based learning mechanism in SyLG that effectively eliminates irrelevant dependency structures. Finally, the Self-adaptive Aggregation Network will fuse the SeSG branch and SyLG branch by cross-attention aggregation mechanism, which considers the complementarity across diverse structures. Our proposed S^2GSL framework makes the following contributions:

- In contrast to leading approaches in complex graph structure learning for ABSA, our proposed SeSG branch introduces constituent trees to decompose the global structure learning into multiple localized substructure learning processes.

- In order to reduce dependence on prior structures, our proposed SyLG branch introduces a learnable method to incorporate syntactic dependencies into latent tree construction.

- Within the two graph learning branch, we propose a Self-adaptive Aggregation Network to facilitate interactions and foster complementary across diverse structures.

- We conduct extensive experiments to study the effectiveness of S^2GSL. Experiments on four benchmarks demonstrate S^2GSL outperforms the baselines. Additionally, the source code and preprocessed datasets used in our work are provided on GitHub[1].

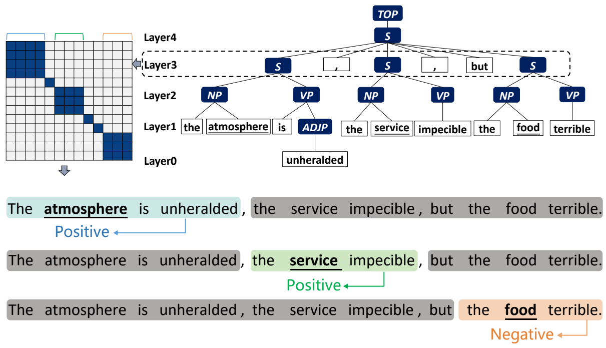


Figure 3: A segment matrix at the third layer of the sample sentence constituent tree divides the sentence into three semantically complete paragraphs.

2 Proposed S^2GSL

The overall architecture of S^2GSL is shown in Figure 4 which is mainly composed of four modules: Context Encoding Module, Segment-aware Semantic Graph Learning(SeSG), Syntax-based Latent Graph Learning(SyLG), and Self-adaptive Aggregation Module. Next, components of S^2GSL will be introduced separately in the rest of the sections.

2.1 Context Encoding Module

Given a sentence of n words $s = \{w_1, w_2, \ldots, w_{\gamma+1} \ldots, w_{\gamma+m} \ldots, w_n\}$, where the aspect $a = \{w_{\gamma+1}, \ldots, w_{\gamma+m}\}$, we use the pre-trained language model BERT (Devlin et al., 2019) as sentence encoder to extract contextual representations. For the BERT encoder, we follow BERT-SPC (Song et al., 2019) to construct a BERT-based sequence x = ([CLS] s [SEP] a [SEP]), if there are multiple aspects in the sentence, we would construct multiple inputs in the format of x. Then the output representation $H^c = \{h^c_1, h^c_2, \ldots, h^c_n\} \in$

[1]https://github.com/ouy7han/S2GSL

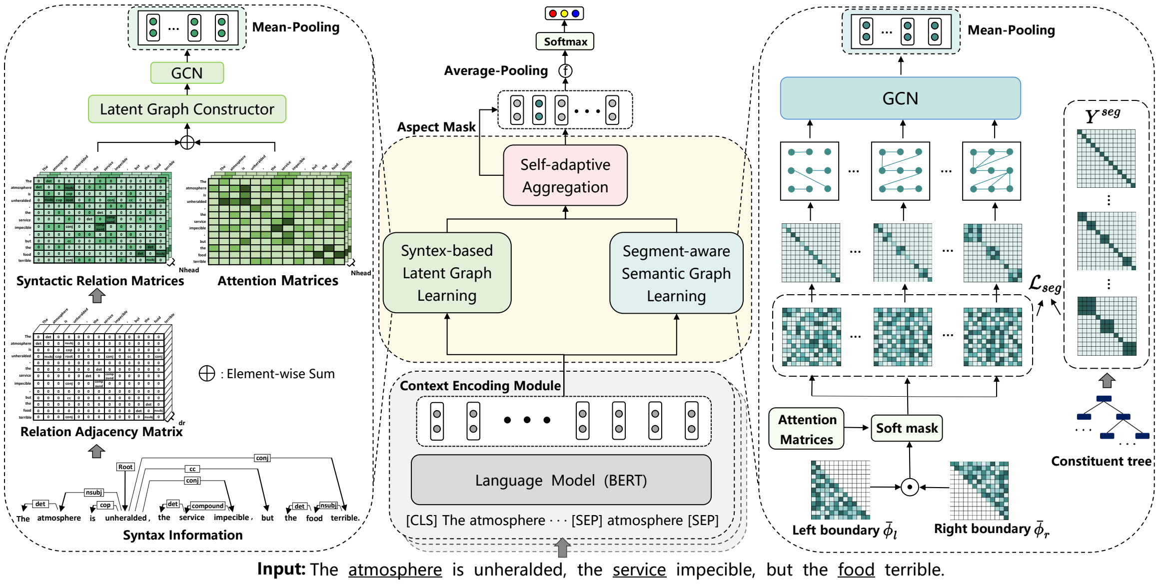


Figure 4: The overall architecture of S[2]GSL, which is composed primarily of SeSG, SyLG, and Self-adaptive Aggregation Modules.

$\mathbb{R}^{n\times d}$ is obtained, where d denotes the dimension of the representation and the c denotes "context".

2.2 Segment-aware Semantic Graph Learning

Since aspects are vulnerable to irrelevant context, inspired by recent work(Nguyen et al., 2020; Shang et al., 2021), we use an end-to-end trainable soft masking dynamic local attention mechanism to construct a SeSG branch aiming to align each aspect and its corresponding opinion.

Attention Segment Masking Matrix We first generate the word-level attention segments for each sentence by training left and right boundary soft masking matrices $\bar{\phi}_l, \bar{\phi}_r \in \mathbb{R}^{n\times n}$, the formulations are calculated as below:

$$\bar{\phi}_l = \text{Softmax}\left(\frac{QW_L^Q(KW_L^K)^T}{\sqrt{d}} \odot \hat{M}\right) \quad (1)$$

$$\bar{\phi}_r = \text{Softmax}\left(\frac{QW_R^Q(KW_R^K)^T}{\sqrt{d}} \odot \hat{M}^T\right) \quad (2)$$

$$\hat{M}_{ij} = \begin{cases} 1, & i \geq j \\ -\infty, & i < j \end{cases} \quad (3)$$

where Q=K=H^c, $\bigodot$ is the element-wise product,and $W_L^Q, W_L^K, W_R^Q, W_R^K \in \mathbb{R}^{d\times d}$ are trainable parameters. Notably, a mask matrix $\hat{M}$ is introduced to ensure that the left boundary position lp and the right boundary position rp generated at position i satisfy $0 \leq lp \leq i \leq rp \leq N$.

The attention segment masking matrix M_s can be obtained by compositing the left and right boundary soft masking matrices $\bar{\phi}_l$ and $\bar{\phi}_r$:

$$M_s = (\bar{\phi}_l L_N) \odot (\bar{\phi}_r L_N^T) \quad (4)$$

where $L_N \in \{0,1\}^{n\times n}$ refers to the upper-triangular matrix.

Then we combine the attention segment masking matrix M_s with the multi-head attention matrices to enable the model to more focus on the semantically relevant contextual information around each word:

$$A^{SeS} = \text{Softmax}\left(\frac{QW^Q(KW^K)^T}{\sqrt{d}} \odot M_s\right) \quad (5)$$

where W^Q, W^K are the trainable parameters, A^{SeS} is a multi-head attention matrix with the number of l, where l corresponds to the number of layers in the constituent tree.

Supervised Constraint In the absence of supervised signal, dynamic local attention may not be able to effectively comprehend the semantically complete segment information around each word, so we further introduce segment-supervised signal to facilitate the learning of dynamic local attention. Specifically, we use the binary cross-entropy loss to represent the distinction between the attention matrix A^{SeS} and the segment-supervised signal Y^{seg}:

$$\mathcal{L}_{seg} = BCE(\sigma(A^{SeS}), Y^{seg}) \quad (6)$$

$$Y_{ij}^{seg} = \begin{cases} 1, & S_{ij} = 1 \\ 0, & else \end{cases} \quad (7)$$

where σ represents the sigmoid function, S_{ij} = 1 indicates the i-th word and the j-th word belong to the same segment, Y^{seg} refers to the segment-supervised signal at each layer of the constituent tree.

To effectively learn the representation of each word, we utilize graph convolutional network (Kipf and Welling, 2017) to extract the segment-aware semantic features H^{SeS} = { h_1^{SeS}, h_2^{SeS}, $\ldots$, h_n^{SeS}}, which is formulated as below:

$$H_l^{SeS} = \sigma(A^{SeS} W_l H_{l-1}^{SeS} + b_l) \tag{8}$$

where H_l^{SeS} represents the l-th GCN output, $H_0^{SeS} = H^c$ is the initial input, σ denotes a non-linear activation function, and W_l and b_l are the trainable parameters.

2.3 Syntax-based Latent Graph Learning

Sharing the same idea with past work(Tang et al., 2022), we also adopt syntactic dependency labels to enhance the latent tree construct. The difference from the past work is that we introduce an attention mechanism in the latent tree to effectively eliminate irrelevant dependency structures and construct a SyLG module.

Syntactically Enhanced Weight Matrix To leverage dependency label information, we first use an off-the-shelf toolkit to obtain dependency information, then we utilize this information to generate a dependency type matrix $R = \{r_{i,j}\}_{n \times n}$, where $r_{i,j}$ represents the types of dependency between x_i and x_j (Tian et al., 2021). Subsequently, we embed each dependency type $r_{i,j}$ into the vector $e_{ij} \in \mathbb{R}^{1 \times dr}$, and finally obtain the relational adjacency matrix $M_R = \{e_{ij} | 1 \leq i \leq n, 1 \leq j \leq n\}$, where e_{ij} refers to the embedding vector of dependency type between the i-th word and the j-th word. If w_i and w_j are not connected, we assign a "0" embedding vector to e_{ij}.

In order to induce a syntax-based latent tree, we need to generate a syntactically enhanced weight matrix. Specifically, we first use multi-head self-attention mechanism to compute a weight matrix A_a. Then, we transform the relation adjacency matrix M_R into a syntactic relation weight matrix A_r through a linear transformation, which has the same number of heads as A_a. Finally, the syntactically enhanced weight matrix $\bar{A}$ can be obtained by summing A_a and A_r:

$$A_a^{\mathrm{k}} = \mathrm{softmax}\left(\frac{QW^Q \times (KW^K)^{\mathrm{T}}}{\sqrt{d}}\right) \tag{9}$$

$$A_r = (W_R^T M_R + b_R) \tag{10}$$

$$\bar{A} = \mathrm{softmax}\left(A_r + A_a\right) \tag{11}$$

where A_a^k is the attention score matrix of the k-th head, $W_R \in \mathbb{R}^{d_r \times N_{head}}$ is the weight matrix for the linear transformation.

Syntax-based Latent Tree Construction Considering $\bar{A}$ as the initial weight matrix, we follow (Zhou et al., 2021) to generate a syntax-based latent tree. We firstly define a variant of the Laplacian matrix for the syntax-based latent tree:

$$\bar{L}_{ij} = \begin{cases} \Phi_i + \sum_{i'=1}^{n} \bar{A}_{i'j} & \text{if } i = j \\ -\bar{A}_{ij} & \textit{otherwise} \end{cases} \tag{12}$$

where $\Phi_i = \exp(W_r h_i^c + b_r)$ is the non-normalized score that the i-th node is selected as the root node, $\bar{L}$ can be used to simplify the computation of weight sums. Therefore, the marginal probability A_{ij}^{SyL} of the syntax-based latent tree can be computed using $\bar{L}_{ij}$:

$$\begin{aligned} A_{ij}^{SyL} &= \left(1 - \delta_{1,j}\right) \bar{A}_{ij} \left[\bar{L}^{-1}\right]_{jj} \\ &\quad - \left(1 - \delta_{i,1}\right) \bar{A}_{ij} \left[\bar{L}^{-1}\right]_{ji} \end{aligned} \tag{13}$$

where δ is Kronecker delta, A^{SyL} can be regarded as the adjacency matrix of the syntax-based latent tree.

We employ a root constraint strategy (Zhou et al., 2021) to direct the root node towards the aspect word:

$$\mathcal{L}_r = -\sum_{i=1}^{N} (t_i \log \boldsymbol{P}_i^r) + (1 - t_i) \log\left(1 - \boldsymbol{P}_i^r\right) \tag{14}$$

where $\boldsymbol{P}_i^r = \Phi_i [\bar{L}^{-1}]_{i1}$ denotes the probability of the i-th word being the root node, and $t_i \in \{0, 1\}$ indicates whether the i-th word is an aspect word.

Similar to the SeSG, we utilize graph convolutional network to extract the syntax-based latent Graph features H^{SyL} = { h_1^{SyL}, h_2^{SyL}, $\ldots$, h_n^{SyL}}, which is formulated as below:

$$H_l^{SyL} = \sigma(A^{SyL} W_l H_{l-1}^{SyL} + b_l) \tag{15}$$

where H_l^{SyL} denotes the l-th layer of GCN output, $H_0^{SyL} = H^c$ is the initial input, W_l and b_l are the trainable parameters.

2.4 Self-adaptive Aggregation Module

Considering the complementarity between SeSG and SyLG, we design a self-adaptive aggregation

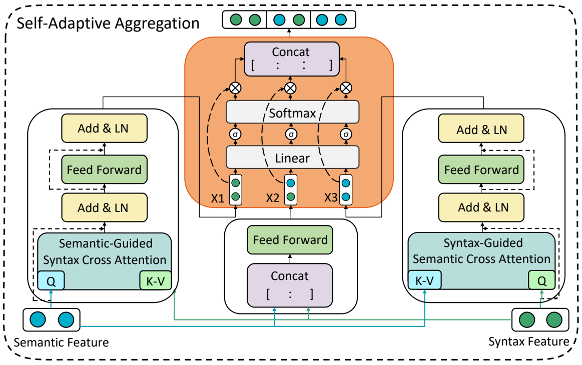


Figure 5: The overall architecture of Self-adaptive Aggregation Module.

module, as shown in Figure 5, to realize their interaction. Specifically, this module consists of three streams, with two of them extended from the traditional Transformer (Vaswani et al., 2017). These two streams can combine information from SeSG and SyLG and ultimately obtain semantic-guided syntax representations H^{SemG} and syntax-guided semantic representations H^{SynG}:

$$O^{Sem} = LN(MH(Q = H^{SeS}, K = V = H^{SyL}) + H^{SeS}) \tag{16}$$

$$H^{SemG} = LN(FFN(O^{Sem}) + O^{Sem}) \tag{17}$$

$$O^{Syn} = LN(MH(Q = H^{SyL}, K = V = H^{SeS}) + H^{SyL}) \tag{18}$$

$$H^{SynG} = LN(FFN(O^{Syn}) + O^{Syn}) \tag{19}$$

where $H^{SemG}, H^{SynG} \in R^{n\times d}$, $MH(\cdot)$ denotes multi-head attention, $LN(\cdot)$ refers to layer normalization, and $FFN(\cdot)$ represents the feed-forward neural network.

To avoid bias towards specific module information, we introduce an additional channel to balance the information between SeSG and SyLG. The specific approach is as follows:

$$H^{Com} = FFN(Concat([H^{SeS}, H^{SyL}])) \tag{20}$$

where $H^{Com} \in R^{n\times d}$, $Concat(\cdot)$ represents the concatenation function.

Considering the different roles of the various module outputs, we assign different weights to these outputs to allow the model to more focus on the important module. Technically, given the input features $X = [X_1, X_2, X_3]$. The weight of each module is calculated by the following equation:

$$a_i = \mathrm{ReLU}(W^T X_i + b) \tag{21}$$

$$\alpha_i = \frac{exp(a_i)}{\sum_{j=1}^{3} exp(a_j)} \tag{22}$$

where $X_1 = H^{SemG}$, $X_2 = H^{SynG}$, $X_3 = H^{Com}$, W_l and b_l are the trainable parameters.The final output feature H^F is generated as follows:

$$H^F = Concat([\alpha_1 H^{SemG}, \alpha_2 H^{SynG}, \alpha_3 H^{Com}]) \tag{23}$$

where $H^F \in R^{n\times d_3}$, with d_3 = 3d.

2.5 Training

We use average pooling at the final aspect nodes of H^F to obtain the aspect representation H^a. Then, the sentiment probability distribution $y_{(s,a)}$ is calculated using a linear layer with a softmax function:

$$y_{(s,a)} = \mathrm{softmax}(W^p H^a + b^p) \tag{24}$$

where (s,a) represents the sentence-aspect pair. Our training objective is to minimize the following objective function:

$$\mathcal{L}(\Theta) = \mathcal{L}_C + \lambda_1 \mathcal{L}_{seg} + \lambda_2 \mathcal{L}_r \tag{25}$$

where Θ denotes all trainable parameters of the model, λ_1 and λ_2 are hyper-parameters, and $\mathcal{L}_C$ is the standard cross-entropy loss function:

$$\mathcal{L}_{\mathrm{C}} = -\sum_{(s,a)\in D}\sum_{c\in C} \log y_{(s,a)} \tag{26}$$

where D contains all sentence-aspect pairs and C is the collection of different sentiment polarities.

3 Experiments

3.1 Datasets

We conduct experiments on four public datasets. The Restaurant and Laptop reviews are from SemEval 2014 Task 4 (Pontiki et al., 2014). The Twitter dataset is a collection of tweets (Dong et al., 2014). The MAMS dataset is consisted of sentences with multiple aspects(Jiang et al., 2019). Each aspect in the sentence is labeled with one of the three sentiment polarities: positive, neutral, and negative. The statistics for the four datasets are shown in Table 1.

Dataset	#Positive		#Negative		#Neutral	
	Train	Test	Train	Test	Train	Test
Laptop	976	337	851	128	455	167
Restaurant	2164	727	807	196	637	196
Twitter	1507	172	1528	169	3016	336
MAMS	3380	400	2764	329	5042	607

Table 1: Statistics for the four experimental datasets.

Model	Laptop		Restaurant		Twitter		MAMS	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
RAM (Chen et al., 2017)	74.49	71.35	80.23	70.80	69.36	67.30	-	
MGAN (Fan et al., 2018)	75.39	72.47	81.25	71.94	72.54	70.81	-	
R-GAT (Wang et al., 2020)	78.21	74.07	86.60	81.35	76.15	74.88	84.52	83.74
KumaGCN (Chen et al., 2020)	<u>81.98</u>	<u>78.81</u>	86.43	80.30	77.89	77.03	-	
ACLT (Zhou et al., 2021)	79.68	75.83	85.71	78.44	75.48	74.51	-	
T-GCN (Tian et al., 2021)	80.88	77.03	86.16	79.95	76.45	75.25	83.38	82.77
DualGCN (Li et al., 2021)	81.80	78.10	87.13	81.16	77.40	76.02	-	
SSEGCN (Zhang et al., 2022)	81.01	77.96	87.31	81.09	77.40	76.02	-	
dotGCN (Chen et al., 2022)	81.03	78.10	86.16	80.49	<u>78.11</u>	77.00	<u>84.95</u>	<u>84.44</u>
MGFN (Tang et al., 2022)	81.83	78.26	<u>87.31</u>	<u>82.37</u>	**78.29**	**77.27**	-	
TF-BERT(dec) (Zhang et al., 2023)	81.49	78.30	86.95	81.43	77.84	76.23	-	
S^{2}GSL(Ours)	**82.46**	**79.07**	**87.31**	**82.84**	77.84	<u>77.11</u>	**85.17**	**84.74**

Table 2: The main experimental results on four public datasets. The best are in bold, and second-best are underlined.

Model	Laptop		Restaurant		Twitter		MAMS	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
S^{2}GSL(Ours)	**82.46**	**79.07**	**87.31**	**82.84**	**77.84**	**77.11**	**85.17**	**84.74**
w/o SyLG	80.41	77.39	86.50	80.22	75.92	74.87	83.71	83.23
w/o SeSG	79.46	76.33	86.10	79.66	76.51	75.23	83.42	82.76
w/o Self-Adaptive Aggregation	80.88	77.07	86.32	80.08	76.07	75.54	83.60	83.05

Table 3: Experimental results of ablation study.

3.2 Implementation Details

The Stanford parser[2] (Manning et al., 2014) is used to obtain syntactic dependencies. Specifically, we use CRF constituency parser (Zhang et al., 2020) to obtain the constituent tree. We use the bert-base-uncase[3] model as our context encoder. The model training is conducted using the Adam optimizer with a learning rate of 2×10^{-5} and L2 regularization of 10^{-5}. The GCN layers of SeSG and SyLG are set to 3. Our model is trained in 20 epochs with a batch size of 16. The hyper-parameters λ_1 and λ_2 for the four datasets are (0.1,0.5), (0.1,0.45), (0.35,0.3) and (0.4,0.75). All experiments are conducted on an NVIDIA 3090 GPU. The model with the highest accuracy or F1 score among all evaluation results is selected as the final model.

3.3 Baselines

We compare our S^{2}GSL with some mainstream and lasted models in ABSA, including **Attention-based methods**: RAM (Chen et al., 2017), MGAN (Fan et al., 2018). **Syntactic-based methods**: R-GAT(Wang et al., 2020),T-GCN(Tian et al., 2021). **Latent-graph methods**: ACLT (Zhou et al., 2021), dotGCN (Chen et al., 2022). **Multi-graph combined methods**: KumaGCN (Chen et al., 2020), DualGCN (Li et al., 2021), SSEGCN(Zhang et al., 2022) MGFN (Tang et al., 2022). **Other method**: TF-BERT (Zhang et al., 2023).

3.4 Overall Performance

All baseline results on four datasets are shown in Table 2, we can find that our S^{2}GSL outperforms all baselines on Laptop, Restaurant, and MAMS datasets. We got the second best on the Twitter dataset, reaching comparable with MGFN. We guess it is because the sentence structure of Twitter dataset is more complex than the other three datasets, basically samples containing only single aspect words, which cannot reflect the advantages of S^{2}GSL. In contrast, the other three datasets, laptop, restaurant, and MAMS contain samples of multiple aspect words. The effect enhancement in Laptop, Restaurant, and MAMS effectively supports that S^{2}GSL gets better sentiment recognition in the case of having multiple aspect words. Additionally, we conduct a parameter comparison between S^{2}GSL and the GCN-based baseline methods. Notably, the number of parameters in S^{2}GSL is comparable (detailed results can be found in A.2). These results demonstrate that S^{2}GSL exhibits superior performance while incurring the same computational overhead.

[2]https://stanfordnlp.github.io/CoreNLP/
[3]https://github.com/huggingface/transformers

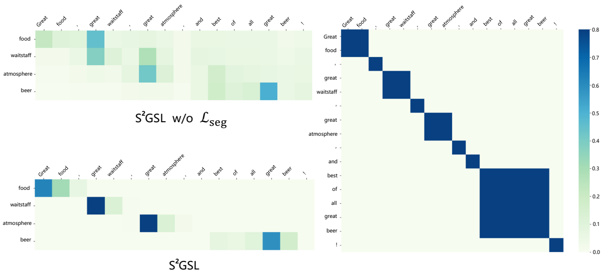


Figure 6: Visualization of attention weights for all aspects (left) and segment-supervised signal (right).

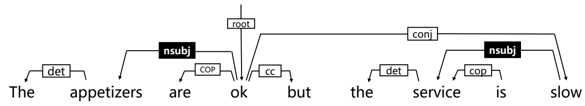
3.5 Ablation Study

We conduct ablation experiments to further investigate the effects of different modules, shown in Table 3. Excluding the syntax-based latent graph learning (w/o SyLG) results in a decrease in the model's performance, which demonstrates the importance of the ability to adaptively capture syntactic relationships between words. We remove the segment-aware semantic graph learning module (w/o SeSG). Compared to the Twitter dataset, the performance of the Restaurant, Laptop and MAMS datasets decreases significantly, which is due to the fact that the sentence structures on these three datasets are more formal and each sentence can be constructed with multiple subordinate clauses. w/o Self-Adaptive Aggregation refers to the SyLG and SeSG modules cannot interact with each other, leading to a drop in performance on all four datasets. The study reveals that both SyLG and SeSG branches are crucial for handling complex sentences in all datasets, as removing either component leads to a noticeable drop in performance. However, compared to the other three datasets, the performance of the SeSG module on the Twitter dataset is not particularly significant, since each sentence on Twitter contains only one aspect word. This difference underscores the adaptability and effectiveness of S^{2}GSL in varying complexity levels across datasets.

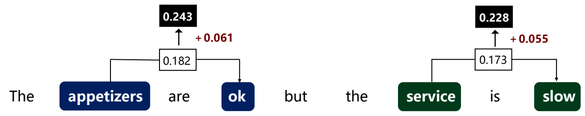
4 Discuss and Analysis

4.1 Effect of Dynamic Local Attention

To demonstrate the effectiveness of supervised dynamic local attention, we visualize the attention weight of all aspects in a sentence, as shown in Figure 6. We can find that in the model without the constraints of segment-supervised signal (S^{2}GSL w/o $\mathcal{L}_{seg}$), the aspect "food" incorrectly assigns a higher attention weight to the opinion "great" of "waitstaff". In contrast, attention weights of each aspect word constrained by the supervised signals are concentrated in its semantically coherent segment and each aspect assigns a higher attention weight to its corresponding opinion, which helps avoid the influence of noisy information.



(a) Dependency tree of an example sentence. Respective aspects and opinions are connected through the "nsubj" label.



(b) The weight change of each aspect word assigned to the corresponding opinion.

Figure 7: Effect of syntactic dependency label.

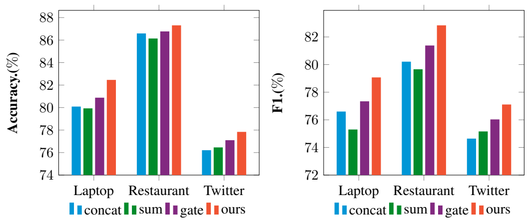


Figure 8: Effects of different fusion strategies.

4.2 Effect of Syntactic Dependency Label

To investigate the validity of syntactic dependency label information, we analyzed the weight change of each aspect word assigned to the corresponding opinion word in the latent tree. Figure 7(b) shows the weight changes of the aspect words "appetizers" and "service" assigned to the corresponding opinion "ok" and "slow". As can be seen from the figure, their respective weights have increased by 0.061 and 0.055. This is because in most cases, each aspect is connected to its corresponding opinion through the same syntactic dependency label, such as the "nsubj" label in Figure 7(a). The above analysis indicates that dependency labels can better capture the relationship between aspects and their corresponding opinions.

4.3 Effects of different fusion strategies

To validate the effectiveness of our proposed self-adaptive aggregation module, we compare it with several typical information fusion strategies: "concat", "sum" and "gate". As shown in Figure 8, we can find that "gate" outperforms better than "concat" and "sum" on all datasets. Furthermore, "con-

Model	Laptop14	Restaurant14
	F1	F1
5-shot ICL (Han et al., 2023)	76.76	81.85
Prompt-setting 1	75.62	79.48
Prompt-setting 2	77.02	81.38
Prompt-setting 3	77.91	82.14
S^2GSL(Ours)	**79.09**	**82.84**

Table 4: The results obtained using ChatGPT and S^2GSL.

cat" performs better than "sum" on the laptop and restaurant, while "sum" performs better on Twitter. These results provide evidence that direct fusion strategies (e.g., concat and sum) are sub-optimal. In contrast, our proposed fusion module achieved the best performance on all datasets, which proves that our fusion module adaptively fuses the respective information in a multi-stream manner, which can fully utilize the complementarities between each stream.

4.4 Experiments With ChatGPT

ChatGPT (OpenAI, 2023a), powered by GPT-3.5 and GPT-4, can achieve significant zero-shot and few-shot in-context learning (ICL) (Brown et al., 2020b) performance on unseen tasks, even without any parameter updates.

In this section, we investigate the performance of ChatGPT on ABSA tasks and its ability for fine-grained understanding of segmental context. We experimented with 3 different prompts and their settings as detailed in A.1. We also compared our results with (Han et al., 2023), who investigated the performance of ChatGPT on various information extraction tasks, including ABSA. All results are shown in table 4. From the experiment, we find that when we prompt ChatGPT with some instructions (single aspect and multi-aspect sentences) can better improve the performance, but its best results are still not as good as our model. This situation suggests that ChatGPT possesses the ability to understand fine-grained segmental information to some extent. Perhaps there are ways to better harness this ability, such as incorporating constituent trees and dynamic local attention mechanisms as described in this paper (refer A.1 for details).

4.5 Impact of Constituent Tree Layer Number

To investigate the impact of different layer numbers of the constituent tree, we evaluate the performance of the model with 2 to 5 constituent tree layers on three different datasets. As shown in Figure 9, the best performance of the model is achieved when the number of layers of the constituent tree is 4. When the layer numbers of the constituent tree are lower than 4, the information from the constituent tree cannot fully cover the entire sentence, resulting in the model not being able to fully learn the complete segment information of a sentence. When the layer numbers are greater than 4, the model will repetitively learn redundant segment information, resulting in a decrease in model performance.

4.6 Case Study

We conduct a case study with a few examples, shown in Table 5. The first sentence only has one aspect word, so all models can easily determine the sentiment polarity of the aspect correctly. For the second comparative type of sentence, there is a certain syntactic dependency between the aspect "hamburger with special sauce" and the aspect "big mac", and the SeSG, which lacks syntactic information, cannot handle this type of sentence well. The last cases contain multiple aspects and opinions. DualGCN and SyLG, which lack segment structure awareness, cannot focus on local information around aspects, resulting in incorrect judgments. Our S^2GSL correctly predicts all samples, indicating that it effectively considers the complementarity between segment structures and syntax correlations of a sentence.

5 Related Work

With the rapid development of ABSA, current research can be broadly divided into three main categories attention-based methods, syntactic-based methods, and multi-graph combined methods.

Attention-based methods Recently, various attention mechanisms have been proposed to implicitly construct the semantic relationships between aspects and their context. (Wang et al., 2016; Tang et al., 2016; Ma et al., 2017; Chen et al., 2017; Gu et al., 2018; Fan et al., 2018; Hu et al., 2019; Tan et al., 2019). For instance, (Wang et al., 2016) proposed an attention-based Long Short-Term Memory (LSTM) network for aspect-based sentiment classification. (Ma et al., 2017) proposed an interactive attention network that can model the connection between the target aspect and the context simultaneously. (Hu et al., 2019) propose orthogonal regularization and sparse regularization so that the attention weights of multiple aspects can focus

Sentences	DualGCN	SyLG	SeSG	S^{2}GSL
1. Much more reasonably [**priced**]$_p$ too!	(P✓)	(P✓)	(P✓)	(P✓)
2. New [**hambuger with special sauce**]$_p$ is ok - at least better than [**big mac**]$_n$!	(P✓,O×)	(P✓,N✓)	(P✓,P×)	(P✓,N✓)
3. Perfectly al dente [**pasta**]$_p$, not drowned in [**sauce**]$_o$ - - generous [**portions**]$_p$.	(P✓,N×,P✓)	(P✓,P×,P✓)	(P✓,O✓,P✓)	(P✓,O✓,P✓)

Table 5: Case study experimental results of four different models.The aspect words are included in [], and p, n, and o represent the true "positive", "negative", and "neutral" sentiment polarities.

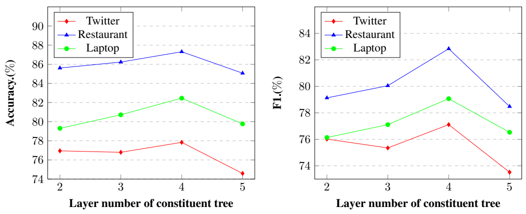


Figure 9: Effect of the number of constituent tree layers.

on different parts of a sentence.

Syntactic-based methods Several studies have explicitly used syntactic knowledge to explicitly build connections between aspects and opinions.(Dong et al., 2014; Zhang et al., 2019; Sun et al., 2019; Huang and Carley, 2019; Zheng et al., 2020; Wang et al., 2020; Zhou et al., 2021). (Wang et al., 2020) reshape the conventional dependency tree with manual rules so that the root node of the dependency tree points to the aspect word. (Zhou et al., 2021) constructed a task-oriented latent tree in an end-to-end fashion.

Multi-graph combined methods With the rapid growth of Graph Convolutional Neural Networks, some studies have explored the combination of different types of graphs for ABSA. For example, (Chen et al., 2020) used a gating mechanism to combine a dependency graph and a latent graph to generate task-oriented representations. (Li et al., 2021) constructed two graph convolutional neural networks using dependency tree and attention mechanism. (Tang et al., 2022) by constructing a latent graph and a semantic graph to effectively capture the interaction between aspects and distant opinions. (Liang et al., 2022) proposes to simultaneously utilize constituent tree and dependency tree of a sentence to model the sentiment relations between each aspect and its context.

However, these methods have primarily relied on the global graph structure learning process, which tends to introduce irrelevant contextual information and syntactic dependency unrelated to specific aspects. Our proposed method of using two graph branches can effectively align each aspect word with its corresponding opinion word.

6 Conclusion

In this paper, we propose an S^{2}GSL model to tackle global structures that will introduce irrelevant contexts and syntactic dependencies during the process of graph structure learning. We propose a SeSG branch that decomposes the ABSA complex graph structure learning problem into multiple local substructure learning processes by utilizing constituent trees. Moreover, we propose a SyLG branch, a more learnable method to introduce syntactic dependencies into latent tree construction. Finally, we devise a Self-adaptive Aggregation Network to realize the interaction between two graph branches, achieving complementarity across diverse structures. Experiments on four benchmarks demonstrate S^{2}GSL outperforms the baselines.

Limitations

S^{2}GSL framework constructs different branches for syntactic and semantic structures which cannot encompass diverse structures in a unified graph modeling process. Therefore, the S^{2}GSL framework further devises an adaptive aggregation to fuse diverse structural information.

Acknowledgements

This research was supported in part by the National Science and Technology Major Project (2021ZD0111501), the National Science Fund for Excellent Young Scholars (62122022), the Natural Science Foundation of China (62206064), the Guangzhou Basic and Applied Basic Research Foundation (2024A04J4384), the major key project of PCL (PCL2021A12), the Guangdong Basic and Applied Basic Research Foundation (2023B1515120020), and the Jihua laboratory scientific project (X210101UZ210).

References

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind

Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020a. Language models are few-shot learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, T. J. Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeff Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020b. Language models are few-shot learners. *ArXiv*, abs/2005.14165.

Chenhua Chen, Zhiyang Teng, Zhongqing Wang, and Yue Zhang. 2022. Discrete opinion tree induction for aspect-based sentiment analysis. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2051–2064, Dublin, Ireland. Association for Computational Linguistics.

Chenhua Chen, Zhiyang Teng, and Yue Zhang. 2020. Inducing target-specific latent structures for aspect sentiment classification. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5596–5607, Online. Association for Computational Linguistics.

Peng Chen, Zhongqian Sun, Lidong Bing, and Wei Yang. 2017. Recurrent attention network on memory for aspect sentiment analysis. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 452–461, Copenhagen, Denmark. Association for Computational Linguistics.

Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam M. Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Benton C. Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier García, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayana Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Díaz, Orhan Firat, Michele Catasta, Jason Wei, Kathleen S. Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2022. Palm: Scaling language modeling with pathways. *J. Mach. Learn. Res.*, 24:240:1–240:113.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. 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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Li Dong, Furu Wei, Chuanqi Tan, Duyu Tang, Ming Zhou, and Ke Xu. 2014. Adaptive recursive neural network for target-dependent Twitter sentiment classification. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 49–54, Baltimore, Maryland. Association for Computational Linguistics.

Feifan Fan, Yansong Feng, and Dongyan Zhao. 2018. Multi-grained attention network for aspect-level sentiment classification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3433–3442, Brussels, Belgium. Association for Computational Linguistics.

Shuqin Gu, Lipeng Zhang, Yuexian Hou, and Yin Song. 2018. A position-aware bidirectional attention network for aspect-level sentiment analysis. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 774–784, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Ridong Han, Tao Peng, Chaohao Yang, Benyou Wang, Lu Liu, and Xiang Wan. 2023. Is information extraction solved by chatgpt? an analysis of performance, evaluation criteria, robustness and errors. *ArXiv*, abs/2305.14450.

Mengting Hu, Shiwan Zhao, Li Zhang, Keke Cai, Zhong Su, Renhong Cheng, and Xiaowei Shen. 2019. CAN: Constrained attention networks for multi-aspect sentiment analysis. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4601–4610, Hong Kong, China. Association for Computational Linguistics.

Binxuan Huang and Kathleen Carley. 2019. Syntax-aware aspect level sentiment classification with graph attention networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language*

Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 5469–5477, Hong Kong, China. Association for Computational Linguistics.

Qingnan Jiang, Lei Chen, Ruifeng Xu, Xiang Ao, and Min Yang. 2019. A challenge dataset and effective models for aspect-based sentiment analysis. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 6280–6285, Hong Kong, China. Association for Computational Linguistics.

Thomas N. Kipf and Max Welling. 2017. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*.

Ruifan Li, Hao Chen, Fangxiang Feng, Zhanyu Ma, Xiaojie Wang, and Eduard Hovy. 2021. Dual graph convolutional networks for aspect-based sentiment analysis. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 6319–6329, Online. Association for Computational Linguistics.

Shuo Liang, Wei Wei, Xian-Ling Mao, Fei Wang, and Zhiyong He. 2022. BiSyn-GAT+: Bi-syntax aware graph attention network for aspect-based sentiment analysis. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 1835–1848, Dublin, Ireland. Association for Computational Linguistics.

Dehong Ma, Sujian Li, Xiaodong Zhang, and Houfeng Wang. 2017. Interactive Attention Networks for Aspect-Level Sentiment Classification. *arXiv e-prints*, page arXiv:1709.00893.

Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, Baltimore, Maryland. Association for Computational Linguistics.

Thanh-Tung Nguyen, Xuan-Phi Nguyen, Shafiq Joty, and Xiaoli Li. 2020. Differentiable window for dynamic local attention. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6589–6599, Online. Association for Computational Linguistics.

OpenAI. 2023a. Introducing chatgpt.

Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. SemEval-2014 task 4: Aspect based sentiment analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.

Xichen Shang, Qianli Ma, Zhenxi Lin, Jiangyue Yan, and Zipeng Chen. 2021. A span-based dynamic local attention model for sequential sentence classification. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 198–203, Online. Association for Computational Linguistics.

Youwei Song, Jiahai Wang, Tao Jiang, Zhiyue Liu, and Yanghui Rao. 2019. Attentional encoder network for targeted sentiment classification. *arXiv preprint arXiv:1902.09314*.

Kai Sun, Richong Zhang, Samuel Mensah, Yongyi Mao, and Xudong Liu. 2019. Aspect-level sentiment analysis via convolution over dependency tree. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 5679–5688, Hong Kong, China. Association for Computational Linguistics.

Xingwei Tan, Yi Cai, and Changxi Zhu. 2019. Recognizing conflict opinions in aspect-level sentiment classification with dual 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)*, pages 3426–3431, Hong Kong, China. Association for Computational Linguistics.

Duyu Tang, Bing Qin, and Ting Liu. 2016. Aspect level sentiment classification with deep memory network. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 214–224, Austin, Texas. Association for Computational Linguistics.

Siyu Tang, Heyan Chai, Ziyi Yao, Ye Ding, Cuiyun Gao, Binxing Fang, and Qing Liao. 2022. Affective knowledge enhanced multiple-graph fusion networks for aspect-based sentiment analysis. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 5352–5362, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Yuanhe Tian, Guimin Chen, and Yan Song. 2021. 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*, pages 2910–2922, Online. Association for Computational Linguistics.

Hugo Touvron, Louis Martin, Kevin R. Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Daniel M. Bikel, Lukas Blecher, Cristian Cantón Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony S. Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel M. Kloumann, A. V. Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, R. Subramanian, Xia Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zhengxu Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models. *ArXiv*, abs/2307.09288.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 6000–6010, Red Hook, NY, USA. Curran Associates Inc.

Kai Wang, Weizhou Shen, Yunyi Yang, Xiaojun Quan, and Rui Wang. 2020. Relational graph attention network for aspect-based sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 3229–3238, Online. Association for Computational Linguistics.

Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. 2016. Attention-based LSTM for aspect-level sentiment classification. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 606–615, Austin, Texas. Association for Computational Linguistics.

Chen Zhang, Qiuchi Li, and Dawei Song. 2019. 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)*, pages 4568–4578, Hong Kong, China. Association for Computational Linguistics.

Mao Zhang, Yongxin Zhu, Zhen Liu, Zhimin Bao, Yunfei Wu, Xing Sun, and Linli Xu. 2023. Span-level aspect-based sentiment analysis via table filling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9273–9284, Toronto, Canada. Association for Computational Linguistics.

Yu Zhang, Houquan Zhou, and Zhenghua Li. 2020. Fast and accurate neural crf constituency parsing. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 4046–4053. International Joint Conferences on Artificial Intelligence Organization. Main track.

Zheng Zhang, Zili Zhou, and Yanna Wang. 2022. SSEGCN: Syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4916–4925, Seattle, United States. Association for Computational Linguistics.

Yaowei Zheng, Richong Zhang, Samuel Mensah, and Yongyi Mao. 2020. Replicate, walk, and stop on syntax: an effective neural network model for aspect-level sentiment classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 9685–9692.

Yuxiang Zhou, Lejian Liao, Yang Gao, Zhanming Jie, and Wei Lu. 2021. To be closer: Learning to link up aspects with opinions. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 3899–3909, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

A Appendix

A.1 Experiments With ChatGPT

The rise of large language models (LLMs) such as GPT-3 (Brown et al., 2020a), PaLM (Chowdhery et al., 2022), Llama (Touvron et al., 2023), etc, has greatly facilitated the rapid development of natural language processing (NLP). ChatGPT (OpenAI, 2023a), powered by GPT-3.5 and GPT-4, can achieve significant zero-shot and few-shot in-context learning (ICL) (Brown et al., 2020b) performance on unseen tasks, even without any parameter updates.

In this section, we conducted exhaustive experiments to investigate the performance of ChatGPT on ABSA tasks and its ability for fine-grained understanding of segmental context.

Experiment setting : The version of ChatGPT we utilized in the experiment is *gpt-3.5-turbo*. To avoid variations in ChatGPT-generated outputs, the *temperature* parameter was set to 0. For the number of response words, the *max_tokens* parameter was set to 512. We experimented with 3 different prompts and their respective settings as shown in Figure10. We take the laptop dataset as example to explain our prompt settings:

- **Prompt-setting 1: Zero-shot.** In this prompt setting, we have only given definition: "Recognize the sentiment polarity for the given aspect term in the given review. Answer from the options [“positive”, “negative”, “neutral”] without any explanation.", then we sent a test review and all aspect words in this review to ChatGPT, and ChatGPT will give the answer.

- **Prompt-setting 2: 5-Shot ICL with single aspect short sentences.** The given definition is same as setting 1. What differs from setting 1 is that we randomly selected 5 single-aspect word sentences from the training set as examples to better leverage ChatGPT’s In-Context Learning capability. Sequentially, we provide a test review and all aspect words, asking ChatGPT to output the sentiment polarity for each aspect word.
- **Prompt-setting 3: 5-Shot ICL with multiple aspects long sentences.** What differs from setting 2 is that the examples chosen from the training set are all multiple aspects long sentences. The purpose of doing this is to explore whether ChatGPT possesses the ability to understand fine-grained segmental context and the strength of this ability.

Result analysis : The results using ChatGPT and our model are shown in table6. Firstly, from the result of prompt-setting 2, we can infer that using a few examples to guide ChatGPT can effectively improve its performance, demonstrating the powerful In-Context-Learning capability of ChatGPT. Secondly, we compared our experimental results with (Han et al., 2023), who investigated the performance of ChatGPT on various information extraction tasks, including ABSA. We found that our model has better performance. Finally, from prompt-setting 3, we can observe that providing ChatGPT with long examples containing multiple aspect words can lead to better performance. However, the results still fall short compared to our model. This situation suggests that ChatGPT possesses the ability to understand fine-grained segmental information to some extent. Perhaps there are ways to better harness this ability, such as incorporating constituent trees and dynamic local attention mechanisms as described in this paper.

A.2 Paramer Comparison

We conduct a parameter comparison between S^{2}GSL and other GCN-based baseline methods, shown in table 7. Notably, the number of parameters in S^{2}GSL is comparable while the two graph learning branch design of S^{2}GSL does lead to an increase in the parameters, it’s important to note that this does not result in an unnecessary escalation in computational costs when compared with the baselines.

Model	Laptop14	Restaurant14
	F1	F1
5-shot ICL (Han et al., 2023)	76.76	81.85
Prompt-setting 1	75.62	79.48
Prompt-setting 2	77.02	81.38
Prompt-setting 3	77.91	82.14
S^{2}GSL(Ours)	**79.09**	**82.84**

Table 6: The results obtained using ChatGPT and S^{2}GSL. All three prompt settings are described in the text.

Model	Parameter Count	Laptop	Restaurant	Twitter
		F1	F1	F1
ACLT(2021)	110M	75.83	78.44	74.51
T-GCN(2021)	113M	77.03	79.95	75.25
DualGCN(2021)	111M	78.10	81.16	76.02
SSEGCN(2022)	110M	77.96	81.09	76.02
S^{2}GSL	**114M**	**79.07**	**82.84**	**77.11**

Table 7: Paramer comparison with other GCN-based baseline methods.

Prompt settings of Laptop dataset

Prompt 1: Zero-shot

" Recognize the sentiment polarity for the given aspect term in the given review. Answer from the options ["positive", "negative", "neutral"] without any explanation.

Review:

" The keyboard has a wonderful nature feel. "

Aspect:

[keyboard]

Answer:

Prompt 2: 5-Shot In-Context-Learning(ICL) with single aspect short sentences

" Recognize the sentiment polarity for the given aspect term in the given review. Answer from the options ["positive", "negative", "neutral"] without any explanation.

Examples:

" Only good thing is the graphics quality. "

Aspect:

[graphics quality]

Answer:

positive

. . . . . . (Four other single aspect short sentences)

Review:

" The keyboard has a wonderful nature feel. "

Aspect:

[keyboard]

Answer:

Prompt 3: 5-Shot ICL with multiple aspects long sentences

" Recognize the sentiment polarity for the given aspect term in the given review. Answer from the options ["positive", "negative", "neutral"] without any explanation.

Examples:

" The video chat is the only thing that is iffy about it but im sure once they unpdate the next version on the mackbook book the quality of it will be better. "

Aspect1: [video chat] Answer: negative

Aspect2: [quality] Answer: positive

. . . . . . (Four other multiple aspects long sentences)

Review:

" The keyboard has a wonderful nature feel. "

Aspect:

[keyboard]

Answer:

Figure 10: Prompt settings of laptop dataset.