Syntactic and Semantic Enhanced Multi-layer Fusion Network for Aspect-based Sentiment Analysis

Kejin Li, Yuanyuan Zhang*, Zhengpeng Li, Jiansheng Wu, Ziteng Wang

***Abstract*—Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment classification task. In this task, graph convolutional networks (GCN) are used to process the syntactic structure and semantic information of sentences, which can effectively capture the expressions of opinion words corresponding to certain aspects. However, using GCN may overlook the overall sentence context, which can lead to incorrect sentiment polarity judgments for aspects. To address the aforementioned issues, we propose the syntactic and semantic enhanced multi-layer fusion network (SSEMFN) model. The model first enhances the ability of GCN to capture the expressions of opinion words corresponding to aspects by using an adjacency matrix that combines syntactic structure and semantic information. Following the GCN, we present our multi-level context-aware module, which improves the model's capacity to comprehend the sentence's overall context. The multi-level context-aware module efficiently captures contextual information at both the local and global levels in sequences. In addition, to prevent loss of feature information due to excessive network depth, residual connections are employed. Experimental results show that our method significantly surpasses the baseline model on three publicly available datasets.**

***Index Terms*—Syntactic and Semantic, Aspect-Based Sentiment Analysis, multi-level context-aware module, Graph Convolutional Networks.**

I. INTRODUCTION

WIth the widespread use of electronic devices, there has been a rise in reviews focusing on specific aspects across various online platforms[1]. To extract targeted opinion information from these reviews[2], aspect-based sentiment analysis (ABSA) has become an area of increasing interest[3]. ABSA, a subtask within sentiment analysis, focuses on identifying the sentiment polarity related to particular aspects in a sentence. Sentiment polarity is categorized as neutral, negative, or positive, with the aspects generally referring to the target entities. For instance, in the

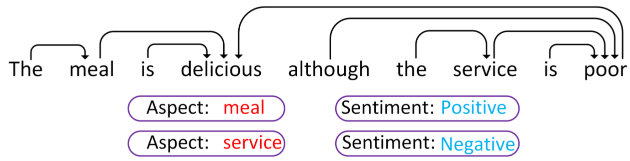

Fig. 1. An example of a dependency tree, with lines indicating the dependencies between words.

example *"The service is decent even when this small place is packed."*, *"service"* and *"place"* are two aspects of the sentence. ABSA can identify the affective polarity of these two aspects, with *"service"* having a positive polarity and *"place"* having a negative polarity.

In earlier research on ABSA, researchers used handcrafted rules and features to train classifiers. Nevertheless, the performance almost reaches a bottleneck due to the tedious nature of manual feature extraction[4], [5]. Recent studies on ABSA have applied graph neural networks (GNN) to the dependency trees of sentences[6], aiming to leverage syntactic structures more effectively. As shown in Figure 1, syntactic dependencies establish links between words within the sentence. In this, a dependency tree is used to construct the dependency graph, which is then treated as the adjacency matrix for the GNN. Among them, Zhang et al.[7] used graph convolutional networks (GCN) to integrate syntactic information from the text. However, GCN treats all neighboring nodes in the graph equally, making it difficult to effectively distinguish the importance of different neighboring nodes. In other words, this means that the model cannot accurately catch the complex relationships between nodes in syntactic information. Huang and Carley et al.[8] used graph attention networks (GAT) to address this issue. Attention mechanism is applied. This enables the GAT network to more effectively consider the importance and relevance of each node to its neighbors, thereby improving the model's capability to identify sentiment polarity. The dependency trees generated by the parser are noisy and unstable. To mitigate dependency parsing errors, Li et al.[9] constructed independent semantic and syntactic graph convolutional networks to learn syntactic and semantic information separately. Along this line, Zhang et al.[10] use GCN to leverage syntactic structure and semantic information to obtain the representations of aspect terms and their corresponding opinion words. However, solely depending on this method to assess the sentiment polarity of aspects may lead to inaccuracies. For instance, in the

Manuscript received November 20, 2024; revised March 20, 2025. The research work was supported by Fundamental Research Project of Higher Education Institutions by Liaoning Provincial Department of Education (LJ222410146057, LJ212410146040), and Graduate Scientific and Technological Innovation Project of University of Science and Technology Liaoning(LKDYC202403).

Kejin Li is a graduate student of University of Science and Technology Liaoning, Anshan, 114051, China. (e-mail: 1710107457@qq.com).

Yuanyuan Zhang* is a associate professor of University of Science and Technology Liaoning, Anshan, 114051, China. (corresponding author, e-mail: yuanyuan@ustl.edu.cn).

Zhengpeng Li is a doctoral student of University of Science and Technology Liaoning, Anshan, 114051, China. (e-mail: lkdlzp0901@163.com).

Jiansheng Wu is a professor of University of Science and Technology Liaoning, Anshan, 114051, China. (e-mail: ssewu@163.com).

Ziteng Wang is a graduate student of University of Science and Technology Liaoning, Anshan, 114051, China. (e-mail: 2957462498@qq.com).

sentence *"With the needs of a professional photographer I generally need to keep up with the best specs."*, the method described above can accurately obtain the representation of the opinion word *"best"* corresponding to the aspect *"specs"*. However, GCN ignores the overall context of the sentence, thus incorrectly judging the emotional polarity of the aspect *"specs"* from neutral to positive. Thus, factoring in the full sentence context is crucial for accurately identifying the sentiment polarity of specific aspects.

This paper presents a novel syntactic and semantic enhanced multi-layer fusion network (SSEMFN) model to address the above issues. First, BERT is utilized as an encoder to gain an encoded representation of the sentence. Next, the self-attention mechanism is used to extract the overall semantic information of the sentence, followed by aspect-aware attention to capture aspect-specific semantics. The matrix of attention scores obtained by combining aspect-aware attention and self-attention was used as the attention neighborhood matrix. The syntactic mask matrix is constructed by considering the varying distances between words in the syntactic dependency tree, resulting in the acquisition of multi-level syntactic information. The attention adjacency matrix and syntactic mask matrix are combined to enhance the ability of GCN to capture the expression of opinion words corresponding to aspects. Subsequently, a multi-level context-aware module is utilized to extract both local and global contextual features in the sequence, improving the model's comprehension of the entire sentence's context. Finally, a residual connection is used to link the input and output of the multi-level context-aware module, resulting in enriched feature representations.

The key contributions of this paper are listed as follows:

1) The SSEMFN model is proposed, which effectively combines syntactic structure and semantic associations while integrating local and global contextual information of sequences.

2)A multi-level context-aware attention module is proposed to improve the model's understanding of the overall sentence context. In addition, residual connectivity is introduced to address the loss of feature information that may occur when the network is too deep.

3)An attention matrix is constructed by combining self-attention with the aspect-aware attention mechanism. The syntactic mask matrix and attention matrix are fused to integrate syntactic structure and semantic information, enhancing GCN's ability to capture opinion word representations related to aspects.

4)Comprehensive tests on benchmark datasets demonstrate that our model outperforms the baseline models.

II. Related Work

ABSA is a classification task. Recently, attention-based networks have been used to semantically model the relationship between aspects and their corresponding context to address issues related to ABSA. Ma et al.[11] employed LSTM to model aspect and context representations independently, and then combined them through interactive attention to generate the final representations. Tan et al.[12] addressed the problem of recognizing conflicting opinions in aspect-level emotion categorization by using two separate attentional mechanisms that focus on positive and negative expressions. In addition, Song et al.[13] used BERT to encode the context and aspects, significantly improving model performance.

Another trend is the use of dependency trees, where syntactic information can effectively link aspects with their corresponding opinion words. GCN utilizes syntactic structures to achieve significant results in the ABSA task. Zhang et al.[7] used GCN to leverage syntactic dependency structures in sentences to obtain aspect features. Liang et al.[14] utilized SenticNet's external sentiment knowledge within the GCN, boosting the model's prediction capabilities. Tian et al.[15] constructed graphs from dependency trees by combining words with dependency types and then applied an attention mechanism to weight the edges in the graphs. These methods have not been successful in effectively combining syntactic structure with semantic correlation. Meanwhile, GCN has not effectively or fully considered the overall context of the sentence when applied to the ABSA task.

III. Model

This section describes in detail our SSEMFN model, whose overall structure is shown in Figure 2. The model is structured into five key components: the input layer, attention layer, syntactic mask layer, GCN layer, and a multi-level context-aware module. The following sections will explain each component in detail.

A. Input Layer

(t,a) is a sentence-aspect pair, where $t = \{w_1, w_2, \ldots, w_l\}$ represents the sentence. $a = \{a_1, a_2, \ldots, a_m\}$ represents an aspect, where a is also a subsequence of t. To better fit the ABSA task, (t,a) is input into BERT in the form "$[CLS] + t + [SEP] + a + [SEP]$" to extract the contextual hidden representation of the sentence aspect pair (t,a), producing the vector $E = \{\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_n\} \in \mathbb{R}^{n \times de}$, with de representing the dimensionality of the word embeddings. After that, the vector E passes through the normalization, dropout and linear layer to obtain the vector representation $H = \{h_1, h_2, \cdots, h_n\} \in \mathbb{R}^{n \times d}$, and d represents the size of the output feature dimension from the linear layer. H contains the subsequence $h_a = \{h_{a_1}, h_{a_2}, \cdots, h_{a_m}\}$ that represents the aspect. H serves as the initial node representation in the model presented in this paper.

B. Attention Layer

Attentional mechanisms are a popular method for capturing interactions between aspects and contextual words[16]. A rich semantic characterization is obtained by combining aspects of perceptual attention and self-attention. As shown in Figure 2, construct p attention adjacency matrices, and then input the obtained adjacency matrices into the syntactic mask layer.

1) Aspect-aware Attention Layer: Aspect-aware attention treats the aspect as a query to capture relevant semantic information, allowing the model to better interpret the sentiment associated with the specific aspect in the sentence. This paper employs p-head aspect-aware attention. The attention score matrix B^l_{asp} derived from the l-th attention head is computed as follows:

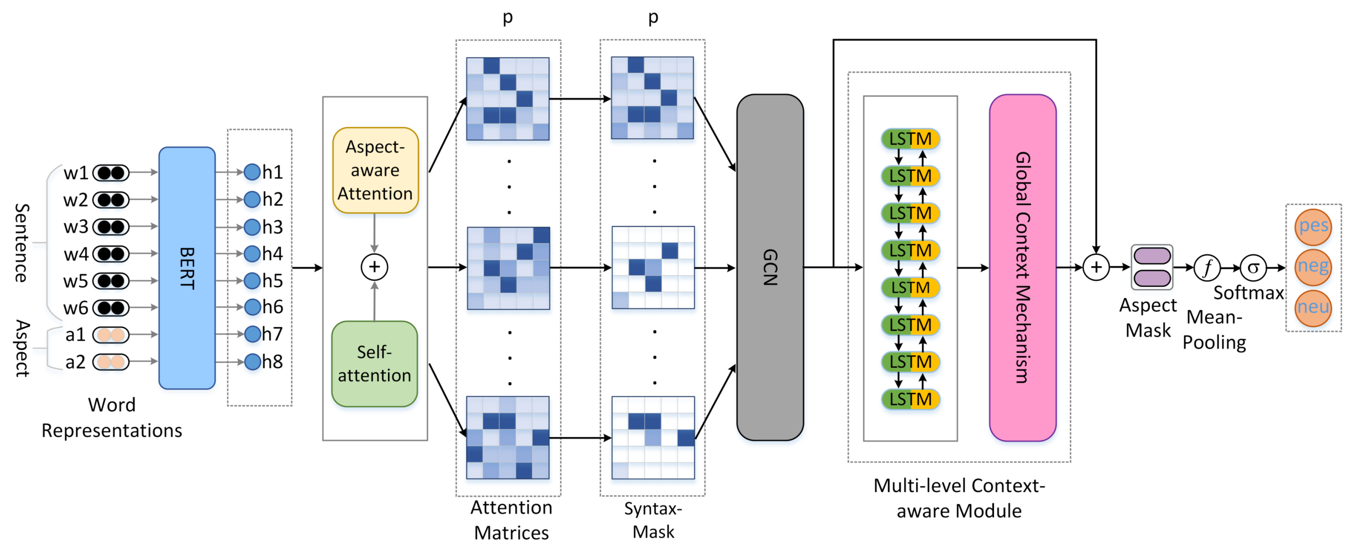

Fig. 2. Overview of SSEMFN model.

$$B_{asp}^{l} = \tanh\left(QW^{a} \times \left(KW^{k}\right)^{T} + b^{a}\right) \tag{1}$$

Here, K is the vector H, and Q is the vector H_a. The aspect representation $H_a \in \mathbb{R}^{n\times d}$ is derived by applying average pooling to h_a and then replicating it n times. b^a is a learnable bias term. $W^k \in \mathbb{R}^{d\times d}$ and $W^a \in \mathbb{R}^{d\times d}$ are learnable weights.

2) Self-Attention Layer: This paper utilizes a p-head self-attention mechanism to grasp the semantic connections between words in a sentence, enabling the extraction of global semantic information. The self-attention score matrix B_{self}^{l} from the l-th self-attention head is computed as follows:

$$B_{self}^{l} = \frac{QW^{\tilde{Q}} \times \left(KW^{\tilde{K}}\right)^{T}}{\sqrt{d}} \tag{2}$$

Here K and Q represent vectors of H, and $W^{\tilde{K}} \in \mathbb{R}^{d\times d}$ and $W^{\tilde{Q}} \in \mathbb{R}^{d\times d}$ are learnable weights. Afterward, B_{asp}^{l} and B_{self}^{l} were added together:

$$B^{l} = B_{asp}^{l} + B_{self}^{l} \tag{3}$$

Where $B^l \in \mathbb{R}^{n\times n}$. Each attention score matrix B^l can be viewed as a fully connected graph, which is then input into the syntactic mask layer.

C. Syntax-Mask Layer

To blend semantic information with syntactic structure information, we derive the syntactic mask matrix based on the varying distances between words in the sentence's syntactic dependency tree. This allows the model to learn local and global syntactic information, which is then combined with the adjacency matrix from the attention layer to enhance GCN performance. Next, we will detail how to construct the syntactic mask matrix.

The syntactic dependency tree is viewed as an undirected graph, with each word considered as a node. $d\,(ti,tj)$ denotes the distance of a path between nodes ti and tj. There may be multiple paths between two nodes, where the shortest path distance is defined as $VT\,(i,j)$:

$$VT\,(i,j) = \min d\,(ti,tj) \tag{4}$$

We set the number of syntactic mask matrices to be the same as the number of attention adjacency matrices in order to achieve masking for each fully connected graph. When the syntactic distance is short, the model captures regional structural information, while a larger syntactic distance allows the model to capture overall structural information. The syntactic mask matrix N^m with a threshold value m is computed as follows:

$$N_{ij}^{m} = \begin{cases} 0, & VT(i,j) \leqslant m \\ -\infty, & \text{otherwise} \end{cases} \tag{5}$$

Where $m \in [1,p]$. Then, each fully connected graph obtained in the attention layers is masked by the corresponding syntactic mask matrix to obtain the adjacency matrix C, where $C \in \mathbb{R}^{p\times n\times n}$.

$$C = softmax\,(B+N) \tag{6}$$

D. GCN Layer

Carry out graph convolution on the previously derived p adjacency matrices C. h^{l-1} represents the input state of the l-th layer, while h^l represents the output state of the l-th layer. The initial state h^0 is the vector H obtained from the input layer, serving as the input to the first layer. In the l-th layer, the nodes update their representations by aggregating the hidden information from their neighboring nodes:

$$h_i^l = \sigma\left(\sum_{j}^{n} C_{ij} W^l h_j^{l-1} + b^l\right) \tag{7}$$

Where σ represents the nonlinear activation function, W^l represents the learnable weight matrix, b^l is the learnable parameter, and h_i^l represents the hidden state of the i-th node in the l-th layer.

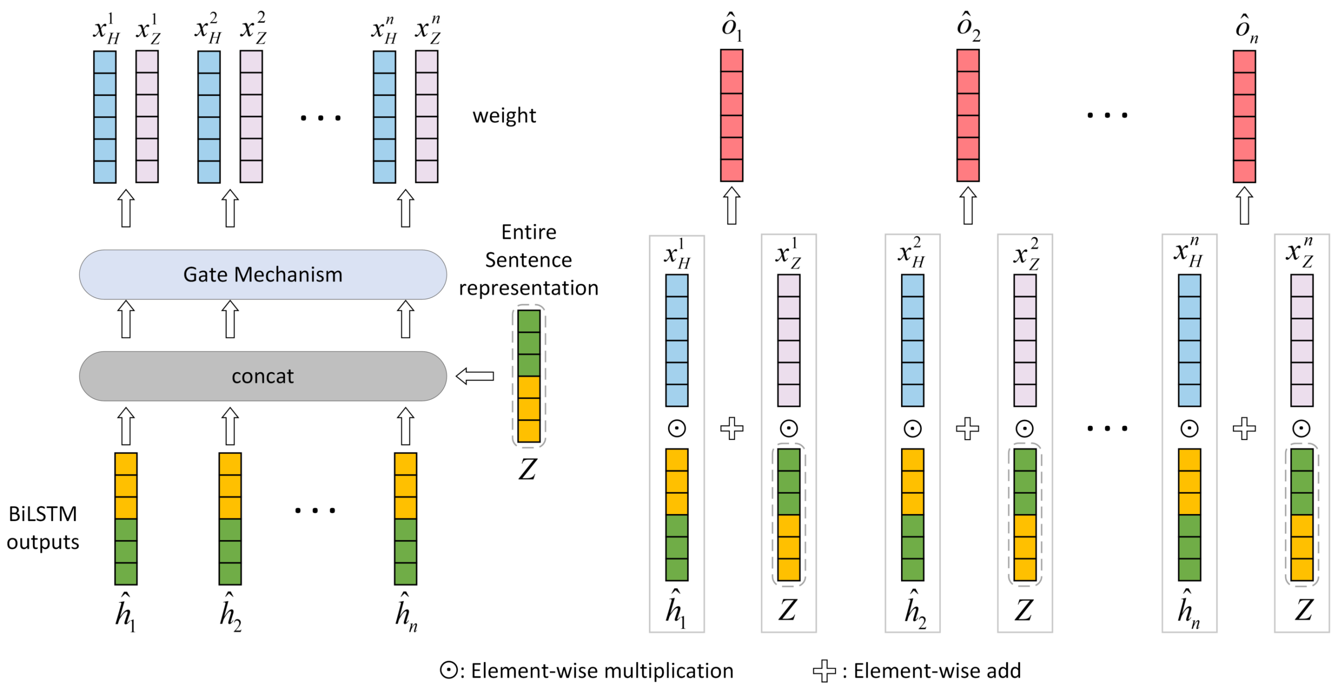

Fig. 3. Diagram of the global context mechanism.

E. Multi-level Context-aware Module

GCN may not adequately capture the overall context of a sentence, which can affect the determination of sentiment polarity for aspects within the sentence. To address this, we propose a multi-level context-aware module, which consists of BiLSTM and a global context mechanism. This module is capable of capturing both local and global contextual information in sequences, thereby enhancing the model's understanding of the overall context.

BiLSTM consists of two LSTMs, one forward processing the input sequence and the other backward processing the input sequence. At each time step, the forward and backward LSTMs generate the hidden states separately, and then they are concatenated together to obtain a complete representation of the current time step. Therefore, BiLSTM can obtain both the preceding and succeeding contextual information at each time step. In this way, BiLSTM is able to effectively capture the local context within these sequences.

$$\vec{h}_i = \overrightarrow{LSTM}\left(\vec{s}_{i-1}, \left[\vec{h}_{i-1} \parallel h_i^l\right]; c_i\right) \tag{8}$$

$$\overleftarrow{h}_i = \overleftarrow{LSTM}\left(\overleftarrow{s}_{i+1}, \left[\overleftarrow{h}_{i+1} \parallel h_i^l\right]; c_i\right) \tag{9}$$

$$\hat{h}_i = \left[\vec{h}_i \parallel \overleftarrow{h}_i\right] \tag{10}$$

Where c_i represents all relevant training parameters of BiLSTM. $\vec{h}_{i-1}$ and $\vec{s}_{i-1}$ denote the hidden state and memorized information of the previous time step respectively, and $\hat{h}_i$ denotes the concatenation of bidirectional hidden states of BiLSTM.

BiLSTM excels at handling local information, which is why we introduced a global context mechanism after BiLSTM. By adding global sentence information to the output of BiLSTM at each time step, the mechanism is able to access the global contextual information of the sequence, thus improving the model's understanding of the overall context. The diagram of the global context mechanism is shown in Figure 3, which will be explained in detail next.

Firstly, in the BiLSTM output, the entire backward sentence information and the entire forward sentence information are distributed across the first and last time steps, respectively. These are concatenated to form $Z = \overleftarrow{h}_1 \parallel \vec{h}_n$, which serves as the entire sentence representation. Each time step output $\hat{h}_i$ of BiLSTM is concatenated with Z to obtain $o_i = Z \parallel \hat{h}_i$, which is fed into the gate mechanism.

In the gate mechanism, a linear layer is used to extract the feature information from o_i, and then a sigmoid function is used to get the weights x_H^i and x_Z^i.

$$x_H^i = sigmoid\left(W_H o_i + b_H\right) \tag{11}$$

$$x_Z^i = sigmoid\left(W_Z o_i + b_Z\right) \tag{12}$$

Where W_H and W_Z are learnable weights, and b_H and b_Z are learnable bias terms. Finally, output $\hat{o}_i$ of the global context is obtained by adding Z and $\hat{h}_i$ with x_Z^i and x_H^i as weights.

$$\hat{o}_i = x_Z^i \odot Z + x_H^i \odot \hat{h}_i \tag{13}$$

Where $\odot$ denotes element-wise multiplication between elements.

In order to prevent the problem that too deep a network may lead to the loss of feature information, residual connections are used to add the global context mechanism and the feature representation obtained by the GCN layer, so as to obtain richer and more comprehensive feature information.

$$g_i = h_i^l + \hat{o}_i \tag{14}$$

$G = \{g_1, g_2, \ldots, g_n\}$ represents the final feature information obtained. Information that is only directly related to the aspect is retained by masking other words. The average pooling operation is then applied to obtain the final aspect information.

$$G_a = f\left(g_{a1}, g_{a2}, \cdots, g_{am}\right) \tag{15}$$

Where $f(\cdot)$ refers to the function for average pooling. G_a is passed through a linear layer, and the softmax function is applied to obtain the probabilities of different sentiment polarities for the aspect.

$$p(b) = softmax\left(W_{\tilde{f}} G_a + b_{\tilde{f}}\right) \tag{16}$$

Here $W_{\tilde{f}}$ and $b_{\tilde{f}}$ represent the learnable weights and bias terms in the linear layer.

F. Model Training

Our model is trained using the cross-entropy loss function, which is defined by the following equation.

$$L(\theta) = -\sum_{(t,a)\in\mathscr{D}} \sum_{c\in\mathscr{C}} \log p(b) \tag{17}$$

In this formula, $\mathscr{D}$ represents all sentence-aspect pairs. θ denotes all trainable parameters in the proposed model, and $\mathscr{C}$ represents the set of different sentiment polarities.

IV. Experiments

A. Datasets and Metrics

The proposed model is tested on three publicly available datasets: Twitter, Laptop, and Restaurant. The Twitter dataset is sourced from tweets collected by Dong et al.[17], while the Laptop and Restaurant datasets are from SemEval2014 Task 4, provided by Pontiki et al.[18]. Each dataset includes three sentiment polarities: neutral, negative, and positive, with each aspect labeled with one of these polarities. Table I displays the statistics for the number of sentiment polarities. As with many ABSA tasks, the evaluation measures used in the experiments are Macro-F1 and Accuracy.

B. Experimental Setup

BERT-base-uncased is used as the encoder for the model. The model uses the Adam optimizer with a learning rate of 2e-5 during training. The batch size of the model is set as 16. This article uses Stanford as a dependency resolver. Other parameter settings used in the experiment are provided in Table II.

TABLE I
Statistics of the three datasets.

Dataset	Neutral		Negative		Positive	
	Train	Test	Train	Test	Train	Test
Twitter	3016	336	1528	169	1507	172
Laptop	455	167	851	128	976	337
Restaurant	637	196	807	196	2164	727

TABLE II
List of hyperparameters adopted for different datasets.

Hyper-parameters	Twitter	Laptop	Restaurant
Epochs	15	10	15
GCN layers	1	2	1
Attention heads	5	5	4

C. Baseline Model

Nine baseline models were selected for comparison. Below is a detailed overview of these models.

BERT[19]: The standard BERT model.

SK-GCN-BERT[20]: To enhance the representation of aspects within sentences, a joint modeling approach is employed, combining syntactic and knowledge information within a single GCN model.

R-GAT-BERT[21]: The method introduces a new dependency tree structure tailored for aspects and utilizes a R-GAT to encode these structures.

DGEDT-BERT[22]: Design a DGEDT network, merging flat representations from Transformers with graph-based features derived from dependency graphs to enhance performance.

BERT4GCN[23]: The syntactic sequence characteristics and syntactic knowledge of the BERT middle layer are combined and integrated with GCN to enhance the coding quality of ABSC tasks.

T-GCN-BERT[15]: First, a graph is constructed using dependency parsing results, integrating word relationships and their dependency types.The edges in the graph are then weighted by applying the attention mechanism. Finally, contextual information from different GCN layers is weighted and combined using an attention layer.

DualGCN-BERT[9]: Two sub-networks, SynGCN and SemGCN, are designed for modeling. Additionally, orthogonal and difference regularizers are proposed to enhance the model's performance.

SSEGCN-BERT[10]: An attention mechanism for aspects, incorporating self-attention, is proposed for obtaining a matrix of attention scores for a sentence, enabling the model to learn the semantic relevance associated with the aspect and the global semantics of the sentence. Syntactic mask matrices of sentences were constructed based on different syntactic distances between words to capture comprehensive syntactic structure information. Combining syntactic structure and semantic information to enhance GCN.

DMGGAT-BERT[24]: The model combines syntactic and semantic information from multi-granularity features of GAT and BERT, which addresses the problem of noise and neglect of sentence-level features in modeling syntactic information. DMGGAT leverages the syntactic and semantic knowledge from BERT to enhance GAT, incorporates an aspect-based attention mechanism to generate sentence-level features, and introduces a multi-granularity gating module to enable the model to capture both aspect and sentence-level features.

D. Main Results

The same dataset was used in order for the experiment to be fair and rational. Table III presents a comparison of the F1 and accuracy scores between the baseline models and our model across three datasets. Our models achieve better performance than the baseline models across all three datasets. Among them, the F1 and accuracy scores of SSEGCN-BERT are significantly better than all other baseline models except for DualGCN-BERT and our SSEMFN. Moreover, on the Twitter dataset, SSEGCN-BERT

TABLE III
COMPARISON OF THREE DATASETS WITH OUR MODEL AND OTHER BASELINE MODELS (%).

Models	Twitter		Laptop		Restaurant	
	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy
BERT	75.18	75.92	76.00	79.91	80.09	85.97
SK-GCN-BERT	73.01	75.00	75.57	79.00	75.19	83.48
R-GAT-BERT	74.88	76.15	74.07	78.21	81.35	86.60
DGEDT-BERT	75.40	77.90	75.60	79.80	80.00	86.30
BERT4GCN	73.76	74.73	73.01	77.49	77.11	84.75
T-GCN-BERT	75.25	76.45	77.03	80.88	79.95	86.16
DualGCN-BERT	76.02	77.40	78.10	81.80	81.16	87.13
SSEGCN-BERT	76.02	77.40	77.96	81.01	81.09	87.31
DMGGAT-BERT	74.56	75.99	77.57	80.78	81.19	87.13
SSEMFN	76.79	78.14	78.96	81.96	82.26	87.49

TABLE IV
ABLATION EXPERIMENT RESULTS (%).

Models	Twitter		Laptop		Restaurant	
	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy
SSEMFN	76.79	78.14	78.96	81.96	82.26	87.49
w/o self-attention	76.73	77.99	77.58	81.01	81.39	87.04
w/o aspect-aware attention	76.49	77.84	77.75	81.01	80.27	85.79
w/o syntactic mask matrix	75.77	77.10	77.20	81.33	81.22	86.86
w/o residual connection	75.94	76.66	74.39	78.96	79.68	86.33
w/o multi-level context-aware module	75.51	76.81	76.45	80.85	79.81	85.70

and DualGCN-BERT have the same F1 and accuracy scores. On the Laptop dataset, the F1 score of SSEGCN-BERT is 0.14% lower than that of DualGCN-BERT, while the accuracy is 0.79% lower. On the Restaurant dataset, the F1 of SSEGCN-BERT is 0.07% lower than DualGCN-BERT, but the accuracy is 0.18% higher than DualGCN-BERT. It is shown that the effectiveness of fusing semantic and syntactic information in ABSA. Our model outperforms the SSEGCN-BERT model, with the F1 improved by 0.77%, 1%, and 1.17% respectively, and accuracy improved by 0.74%, 0.95%, and 0.18% respectively, on Twitter, Laptop and Restaurant datasets. It is also superior to the DualGCN-BERT model, with the F1 improved by 0.77%, 0.86%, and 1.1% respectively, and the accuracy improved by 0.74%, 0.16%, and 0.36% respectively, on the three datasets. This demonstrates that our SSEMFN model strengthens GCN's capability to capture the expressions of opinion words associated with aspects by creating an adjacency matrix that incorporates both syntactic and semantic information. The multi-level context-aware module, by leveraging local and global contextual information, enables a more complete understanding of the sentence's overall context, leading to a more precise judgment of the sentiment polarity of the aspect.

E. Ablation Experiments

To assess the contribution of each module in the SSEMFN model, ablation experiments are conducted on three datasets. In Table IV, where our SSEMFN is the baseline model and w/o means without, such as w/o self-attention indicating the model without self-attention. Without self-attention, the F1 of the model drops by 0.06%, 1.38%, and 0.87% respectively, and accuracy drops by 0.15%, 0.95%, and 0.45% respectively. This suggests that the model lost the sentence's global semantic information, resulting in a noticeable drop in performance. Without aspect-aware attention, the F1 of the model decreased by 0.3%, 1.21%, and 1.99%, respectively, and the accuracy decreased by 0.3%, 0.95%, and 1.7%. This demonstrates the importance of aspect-aware attention in capturing key local semantic connections the context words and the aspect. After removing the syntactic mask matrix, the model's F1 dropped by 1.02%, 1.76%, and 1.04%, while its accuracy scores decreased by 1.04%, 0.63%, and 0.63%, respectively. This indicates that the syntactic mask matrix helps the GCN fully learn syntactic structure information in the dependency tree. After removing the residual connections, the model's F1 decreased by 0.85%, 4.57%, and 2.58%, respectively, and the accuracy decreased by 1.48%, 3%, and 1.16%, respectively. This indicates that residual connections prevent the problem of feature information loss due to too deep a network. When multi-level context-aware module are removed, the model's F1 decreases by 1.28%, 2.51%, and 2.45%, respectively, and the accuracy decreases by 1.33%, 1.11%, and 1.79%, respectively. This shows that the multi-level context-aware module effectively captures both local and global contextual information of the sequence, improving the model's overall understanding of the sentence's context. In summary, the performance of the model declines after removing each module, demonstrating that each module is crucial to our model.

F. The performance is influenced by the GCN layers' number

Experiments were conducted on the Laptop and Restaurant datasets to analyze how different numbers of GCN layers affect the model's performance. The number of GCN layers was tested from 1 to 6, with other parameters held constant. In Figure 4, the 2-layer GCN model achieves the best performance on the Laptop dataset. In Figure 5, the 1-layer GCN model performs best when applied to the Restaurant dataset. The results suggest that increasing the number of

TABLE V
CASE STUDY RESULTS.

Aspect	Sentence	SSEGCN-BERT	SSEMFN	Label
{Boot, time}	**Boot time** is super fast, around anywhere from 35 seconds to 1 minute.	{P}	{P}	{P}
{tech support}	**tech support** would not fix the problem unless I bought your plan for $ 150 plus.	{N}	{N}	{N}
{build, durability}	Strong **build** though which really adds to its **durability**.	{P, P}	{P, P}	{P, P}
{SSD, 16Gb RAM}	I 've installed to it additional **SSD** and **16Gb RAM**.	{O, O}	{O, O}	{O, O}
{USB3 Peripherals, ThunderBolt}	**USB3 Peripherals** are noticably less expensive than the **ThunderBolt** ones .	{P, N}	{P, N}	{P, N}
{disk drive}	It 's ok but does n't have a **disk drive** which I did n't know until after I bought it.	{N}	{O}	{O}
{specs}	With the needs of a professional photographer I generally need to keep up with the best **specs**.	{P}	{O}	{O}
{size}	The smaller **size** was a bonus because of space restrictions .	{N}	{P}	{P}
{functionality}	You just can not beat the **functionality** of an Apple device .	{N}	{P}	{P}
{lunch, dinners}	How pretentious and inappropriate for MJ Grill to claim that it provides power **lunch** and **dinners** !	{P, P}	{N, N}	{N, N}
{staff}	The **staff** should be a bit more friendly .	{P}	{N}	{N}
{sushi}	The **sushi** is cut in blocks bigger than my cell phone .	{P}	{N}	{N}

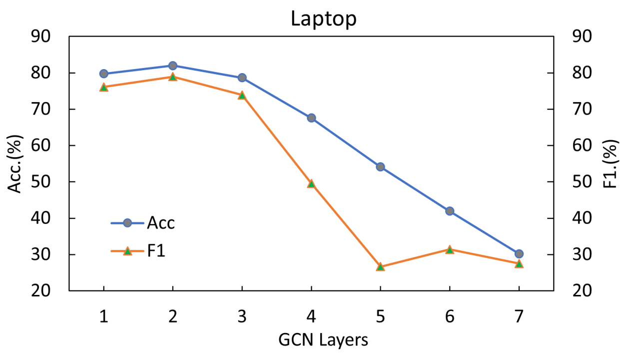

Fig. 4. SSEMFN layer count impact on Laptop dataset.

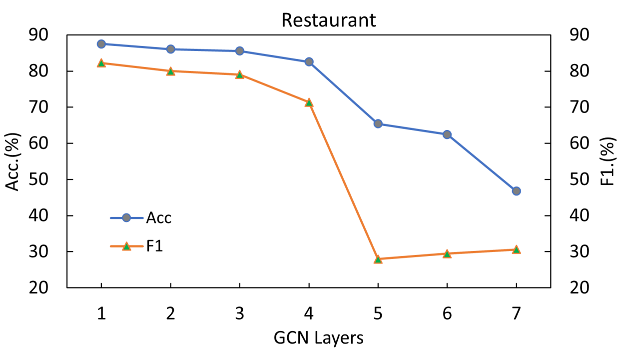

Fig. 5. SSEMFN layer count impact on Restaurant dataset.

GCN layers leads to poorer performance on both datasets. This is because, as the number of GCN layers grows, the node information becomes overly smooth through multiple aggregations, losing the details and distinctions from the original graph structure. Additionally, the repeated use of neighboring node information during aggregation results in node representations containing a large amount of redundant information.

G. Case Study

The case study results of the SSEGCN-BERT model and our proposed model on some sample sentences are shown in Table V. The table contains the predictions of the two models for these sample sentences and the corresponding true labels, with the symbols O, P, and N representing neutral, positive, and negative emotions, respectively. The sentence aspects are in bold. The table shows that both models correctly predict the sentiment polarity of the aspects in the first five sample sentences. This is because both models effectively combine syntactic structure and semantic information to enhance GCN's ability to capture the expressions of opinion words corresponding to aspects. In the sixth sample sentence, not only was our SSEMFN model also able to capture the expression of the opinion word *"n't"* corresponding to the aspect *"disk drive,"* but the model also considered the overall context of the sentence, leading to the conclusion that the overall sentiment should be neutral. As a result, it successfully judged the emotion polarity of the *"disk drive"* as neutral rather than negative. Similarly, in the seventh sample sentence, we also successfully determined the emotion polarity of the aspect *"specs"* to be neutral. Because our SSEMFN model might not only have captured the expression of the opinion word *"best"* corresponding to the aspect, but also considered the overall context of the sentence. We also correctly determined the emotion polarity of the aspects in the last five sentences. However, in the last seven sample sentences, the SSEGCN-BERT model only considered the expression of opinion words corresponding to the aspects without taking the overall context of the sentence into account, leading to errors in predicting the emotion polarity of the aspects. Our SSEMFN model incorporates syntactic structure and semantic information to enhance the ability of GCNs to obtain aspectual word correspondences for opinion word expressions. It also captures both local and global contextual information of sequences through a multi-level context-aware module, enabling an overall contextualization of the sentence to be taken into account by the model. Thereby, accurate predictions of the emotional polarity of aspects can be made.

V. CONCLUSION

We propose an SSEMFN model to address ABSA tasks. Self-attention, combined with aspect-aware attention, forms the self-attention layer to learn global semantic information as well as aspect-related semantic information.

Syntactic mask matrices of sentences are constructed using different syntactic distances to capture syntactic structure information. The combination of the syntactic mask matrix with the attention score matrix, fusing syntactic and semantic information, enhances the ability of the GCN to obtain aspects corresponding to opinion word expressions. To better account for the overall context of sentences, this paper proposes a multi-level context-aware module, which is used after the GCN to capture the multi-level contextual information of the sequence. Residual connections are used to avoid feature information loss from overly deep networks. Tests conducted on three publicly accessible datasets validate the performance of our mode

REFERENCES

[1] D. T. N. El-Deen, R. S. El-Sayed, A. M. Hussein, and M. S. Zaki, "Multi-label classification for sentiment analysis using cbga hybrid deep learning model," *Engineering Letters*, vol. 32, no. 2, pp. 340 – 349, 2024.

[2] C. Henriquez and G. Sanchez-Torres, "Aspect extraction for opinion mining with a semantic model," *Engineering Letters*, vol. 29, no. 1, pp. 61 – 67, 2021.

[3] A. Nazir, Y. Rao, L. Wu, and L. Sun, "Issues and challenges of aspect-based sentiment analysis: A comprehensive survey," *IEEE Transactions on Affective Computing*, vol. 13, no. 2, pp. 845 – 863, 2022.

[4] Erlin, I. Suliani, H. Asnal, L. Suryati, and R. Efendi, "Sentiment analysis for abolition of national exams in indonesia using support vector machine," *Engineering Letters*, vol. 30, no. 4, pp. 1342 – 1352, 2022.

[5] S. Li, Z. Li, J. Wu, J. Miao, Y. Bai, X. Yu, and K. Li, "Attention feature fusion graph convolutional network for target-oriented opinion words extraction," *Engineering Letters*, vol. 31, no. 3, pp. 1273 – 1280, 2023.

[6] N. Isa, M. Puteh, and R. M. H. R. Kamarudin, "Sentiment classification of malay newspaper using immune network (scin)," in *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2013*, vol. 3 LNECS, London, United kingdom, 2013, pp. 1543 – 1548.

[7] C. Zhang, Q. Li, and D. Song, "Aspect-based sentiment classification with aspect-specific graph convolutional networks," in *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Hong Kong, China, 2019, pp. 4568 – 4578.

[8] B. Huang and K. M. Carley, "Syntax-aware aspect level sentiment classification with graph attention networks," in *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Hong Kong, China, 2019, pp. 5469 – 5477.

[9] R. Li, H. Chen, F. Feng, Z. Ma, X. Wang, and E. Hovy, "Dual graph convolutional networks for aspect-based sentiment analysis," in *ACL-IJCNLP 2021 - 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Virtual, Online, 2021, pp. 6319 – 6329.

[10] Z. Zhang, Z. Zhou, and Y. Wang, "Ssegcn: Syntactic and semantic enhanced graph convolutional network for aspect-based sentiment analysis," in *NAACL 2022 - 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, Seattle, WA, United states, 2022, pp. 4916 – 4925.

[11] D. Ma, S. Li, X. Zhang, and H. Wang, "Interactive attention networks for aspect-level sentiment classification," in *IJCAI International Joint Conference on Artificial Intelligence*, vol. 0, Melbourne, VIC, Australia, 2017, pp. 4068 – 4074.

[12] X. Tan, Y. Cai, and C. Zhu, "Recognizing conflict opinions in aspect-level sentiment classification with dual attention networks," in *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*, Hong Kong, China, 2019, pp. 3426 – 3431.

[13] Y. Song, J. Wang, T. Jiang, Z. Liu, and Y. Rao, "Targeted sentiment classification with attentional encoder network," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11730 LNCS, Munich, Germany, 2019, pp. 93 – 103.

[14] B. Liang, H. Su, L. Gui, E. Cambria, and R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks," *Knowledge-Based Systems*, vol. 235, 2022.

[15] Y. Tian, G. Chen, and Y. Song, "Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble," in *NAACL-HLT 2021 - 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, Virtual, Online, 2021, pp. 2910 – 2922.

[16] F. Fan, Y. Feng, and D. Zhao, "Multi-grained attention network for aspect-level sentiment classification," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, Brussels, Belgium, 2018, pp. 3433 – 3442.

[17] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu, "Adaptive recursive neural network for target-dependent twitter sentiment classification," in *52nd Annual Meeting of the Association for Computational Linguistics, ACL 2014 - Proceedings of the Conference*, vol. 2, Baltimore, MD, United states, 2014, pp. 49 – 54.

[18] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, "Semeval-2014 task 4: Aspect based sentiment analysis," in *8th International Workshop on Semantic Evaluation, SemEval 2014 - co-located with the 25th International Conference on Computational Linguistics, COLING 2014, Proceedings*, Dublin, Ireland, 2014, pp. 27 – 35.

[19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," in *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, vol. 1, Minneapolis, MN, United states, 2019, pp. 4171 – 4186.

[20] J. Zhou, J. X. Huang, Q. V. Hu, and L. He, "Sk-gcn: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification," *Knowledge-Based Systems*, vol. 205, 2020.

[21] K. Wang, W. Shen, Y. Yang, X. Quan, and R. Wang, "Relational graph attention network for aspect-based sentiment analysis," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Virtual, Online, United states, 2020, pp. 3229 – 3238.

[22] H. Tang, D. Ji, C. Li, and Q. Zhou, "Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Virtual, Online, United states, 2020, pp. 6578 – 6588.

[23] Z. Xiao, J. Wu, Q. Chen, and C. Deng, "Bert4gcn: Using bert intermediate layers to augment gcn for aspect-based sentiment classification," in *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, Virtual, Punta Cana, Dominican republic, 2021, pp. 9193 – 9200.

[24] Y. Wang, N. Yang, D. Miao, and Q. Chen, "Dual-channel and multi-granularity gated graph attention network for aspect-based sentiment analysis," *Applied Intelligence*, vol. 53, no. 11, pp. 13 145 – 13 157, 2023.