

## Article

# Dual-Channel Edge-Featured Graph Attention Networks for Aspect-Based Sentiment Analysis

Junwen Lu <sup>1</sup>, Lihui Shi <sup>1,\*</sup>, Guanfeng Liu <sup>2</sup> and Xinrong Zhan <sup>1</sup>

<sup>1</sup> School of Computer and Information and Engineering, Xiamen University of Technology, Xiamen 361024, China

<sup>2</sup> School of Computing, Macquarie University, Macquarie Park, NSW 2109, Australia

\* Correspondence: lhshi6@163.com

**Abstract:** The goal of aspect-based sentiment analysis (ABSA) is to identify the sentiment polarity of specific aspects in a context. Recently, graph neural networks have employed dependent tree syntactic information to assess the link between aspects and contextual words; nevertheless, most of this research has neglected phrases that are insensitive to syntactic analysis and the effect between various aspects in a sentence. In this paper, we propose a dual-channel edge-featured graph attention networks model (AS-EGAT), which builds an aspect syntactic graph by enhancing the contextual syntactic dependency representation of key aspect words and the mutual affective relationship between various aspects in the context and builds a semantic graph through the self-attention mechanism. We use the edge features as a significant factor to determine the weight coefficient of the attention mechanism to efficiently mine the edge features of the graph attention networks model (GAT). As a result, the model can connect important sentiment features of related aspects when dealing with aspects that lack obvious sentiment expressions, pay close attention to important word aspects when dealing with multiple-word aspects, and extract sentiment features from sentences that are not sensitive to syntactic dependency trees by looking at semantic features. Experimental results show that our proposed AS-EGAT model is superior to the current state-of-the-art baselines. Compared with the baseline models of LAP14, REST15, REST16, MAMS, T-shirt, and Television datasets, the accuracy of our AS-EGAT model increased by 0.76%, 0.29%, 0.05%, 0.15%, 0.22%, and 0.38%, respectively. The macro-f1 score increased by 1.16%, 1.16%, 1.23%, 0.37%, 0.53%, and 1.93% respectively.

**Keywords:** graph attention networks; ABSA; self-attention; dependent tree; semantic features



**Citation:** Lu, J.; Shi, L.; Liu, G.; Zhan, X. Dual-Channel Edge-Featured Graph Attention Networks for Aspect-Based Sentiment Analysis. *Electronics* **2023**, *12*, 624. <https://doi.org/10.3390/electronics12030624>

Academic Editors: Juan M. Corchado, Byung-Gyu Kim, Carlos A. Iglesias, In Lee, Fuji Ren and Rashid Mehmood

Received: 2 December 2022

Revised: 22 January 2023

Accepted: 23 January 2023

Published: 26 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

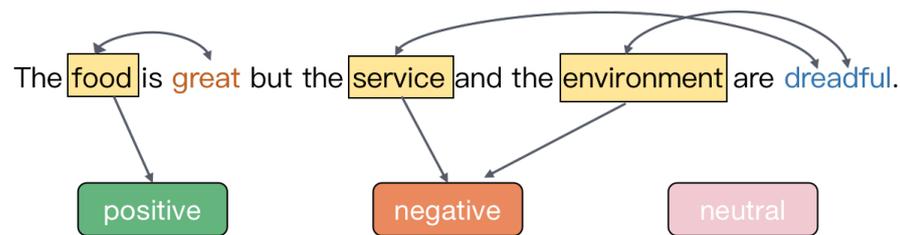
Aspect-based sentiment analysis (ABSA) is an important task in natural language processing (NLP). ABSA is a fine-grained sentiment analysis task that is used to determine the sentiment polarity of individual components of a text (such as positive, negative, or neutral). As illustrated in Figure 1, the sentiment polarity of the aspect “food” is positive, but the sentiment polarity of the aspect “services” and “environment” is negative. As a result, as compared to traditional text sentiment analysis, ABSA can mine users’ more delicate sentiment expressions, analyze users’ distinct sentiment viewpoints on diverse goals, and provide decision-makers with more precise decision help.

On the ABSA task, the early is mainly based on artificial feature engineering, artificial design, and feature selection, such as sentiment dictionary, dependency information, etc., using the maximum entropy, support vector machine, and other traditional classifiers for sentiment classification. Although good performance has been achieved, the shortcomings of poor generalization ability and high labor costs are gradually exposed.

In recent years, due to the rise of neural networks, convolutional neural network (CNN) and recurrent neural network (RNN) have been used to solve problems in ABSA.

For instance, FAN [1] proposed CMN; Li [2] proposed TNet; Xue [3] proposed the GCAE model. Huang [4] proposed PF-CNN and PG-CNN. The TD-LSTM and TC-LSTM models were proposed by Tang [5]. The GRNN model was suggested by Zhang [6]. The H-LSTM model was proposed by Ruder [7].

Many classification tasks, including ABSA, have been completed using the pre-trained language model BERT [8]. As an illustration, Xu [9] utilized a different corpus to post-train BERT and showed how well it performed aspect extraction and ABSA. By creating supplementary sentences, Sun [10] transformed ABSA into a sentence pair classification task.



**Figure 1.** Examples of ABSA tasks. Each aspect is classified into a corresponding sentiment polarity.

With the development of the attention mechanism, the attention mechanism has been widely used in this task, as it can make the model focus on a given aspect, to make a more accurate judgment of sentiment polarity. The ATAE-LSTM model was suggested by Wang [11], the AB-LSTM model by Yang [12], and the BILSTM-ATT model by Liu [13]. The AOA model was proposed by Huang [14] and the IAN model by Ma [15].

Because of their intrinsic capacity in specific aspects and semantic alignment of contextual words, attention mechanisms and CNN are commonly utilized in aspect sentiment analysis. However, because these models lack the mechanism to account for relevant syntactic constraints and distant word dependencies, they may recognize grammatically irrelevant contextual words as signals to judgment sentiments. Zhang [16] presented the ASGCN model to overcome this problem; Hou [17] proposed the SA-GCN model, and Xiao [18] proposed the AEGCN model.

The syntactic structure of sentences is used by graph convolution networks (GCNs) [19] and graph attention networks (GATs) [20] based on dependency trees. This is because syntactic dependence can build the relationship between words in a sentence, but sentences with poor syntactic dependence sensitivity cannot gain appropriate sentiment qualities via the dependency tree. Furthermore, existing ABSA methods typically embed aspect information into sentence representation to learn relevant sentiment features of specific aspects, but this method makes obtaining valuable sentiment features for aspects lacking obvious sentiment features difficult.

We have two issues that need to be resolved, as was previously discussed. (1) How to extract valuable sentiment features from aspects without obvious sentiment expression; (2) How to extract precise Sentiment features from sentences with low syntactic sensitivity. In this paper, we provide a fresh approach to the issues raised above. For the first issue, we create the aspect parsing graph by enhancing the context-dependent encoding of important aspect words and utilizing the emotional reciprocity between various context-related aspects. To solve the second issue, we create the semantic graph by applying the self-attention mechanism to extract the semantic link between words.

Based on this, we present a dual-channel edge-featured graph attention networks model (AS-EGAT). We improve the GAT model's original attention mechanism by making the edge features a significant factor in the calculation of the weight coefficients by the attention mechanism to efficiently mine the edge features of the GAT model, which uses aspect syntactic graphs and semantic graphs as adjacency matrices. The following is a summary of our paper's significant contributions:

- We create aspect syntactic graphs, which are strengthened by the mutuality of various aspects in the context and particular aspect words, as well as semantic graphs,

which use self-attention to determine the semantic connections between each word in a phrase;

- We propose a dual-channel edge-featured graph attention networks model (AS-EGAT), which learns aspect sentiment features by modeling aspect features and semantic features in the context and fully mining the edge features;
- We conduct extensive experiments on six benchmark datasets. The experimental results show the effectiveness of the AS-EGAT model in ABSA.

## 2. Related Work

Many people are now employing CNN to solve the ABSA task because it has lately developed rapidly. A convolutional memory network (CMN) that incorporates the attention mechanism and that can simultaneously capture single words and multi-word expressions in sentences was proposed by Fan [1]. After noting the deficiencies of the attention mechanism and the reasons why CNN is poor at categorizing tasks, Li [2] proposed a new classification model, TNet, to assist the CNN feature extractor to identify sentiment features more precisely. This model utilized an approach strategy and scaled the convolutional layer's input by using the location correlation between words and aspect words. The earliest research employed LSTM and attention mechanisms to predict the sentiment polarity of sentences. According to Xue [3], this kind of model is extremely sophisticated and requires a lot of training time. Based on this, they suggest the GCAE, a convolutional neural network model with a gate mechanism that is more effective, whose gate units can operate independently, and whose computation can be easily parallelized during training. Convolutional neural network parameterized filter (PF-CNN) and parameterized gated Convolutional Neural network (PG-CNN) were proposed by Huang [4]. They included aspect word information into CNN by using the parameterized filter gate structure.

The majority of cutting-edge techniques in numerous ABSA problems are built on RNNs for modeling because of their strong sequence learning capabilities. To further develop the LSTM, Tang [5] presented the target-dependent long-term memory neural network (TD-LSTM) and the target-association-based long-term memory neural network (TC-LSTM). The model treats a given target as a feature for ABSA and links it to contextual features. To capture sentence syntax and semantics as well as the interaction between aspectual words and nearby contextual words via Bi-RNN, Zhang [6] created two gated neural network models (GRNNs) (bidirectional RNN). By more effectively propagating the gradient, gated neural networks have been demonstrated to lower the deviation of standard recurrent neural networks at the end of the sequence. Ruder proposed a comment hierarchy model (H-LSTM) [7] that uses hierarchical bidirectional long and short-term memory neural networks to learn intra-sentence and inter-sentence relationships. The model integrates word input into the bidirectional LSTM at the sentence level, and the forward and backward LSTM's final states are connected by aspect embedding and are fed back into the bidirectional LSTM at the comment level.

The attention mechanism's core goal is to select the information that is more critical to the current task objective from a large amount of information and then use the attention mechanism to obtain the description of a specific aspect category in the text, to make a more accurate polarity judgment of sentiment, which is consistent with the goal of aspect level sentiment classification. Wang [11] was the first to introduce an attention mechanism to this task, proposing an attention-based LSTM model using target embedding (ATAE-LSTM), where an attention vector is set for each target on top of an LSTM network, forcing the model to attend to the important parts of the sentence. This method was shown to be an effective way to strengthen the neural model to attend to the relevant parts of the sentence. Two attention-based bidirectional LSTM were also suggested by Yang [12] to enhance classification performance. Liu [13] expanded the attention model by differentiating between the left and right contexts of a particular target. Huang [14] proposed an attentional overattentive neural network (AOA) that jointly learns the representation of aspects and sentences through AOA modules and automatically focuses on the important parts of the

sentences in a study based on LSTM with a joint approach to model aspects and sentences. Following that, some interactive attention mechanisms were used on ABSA tasks. Ma [15] proposed an interactive attention network (IAN) that considers both target attention and contextual attention, employing two attention networks to detect important words from the target description and important words from the entire context.

Because of their intrinsic capacity to match aspect-specific and contextual word meanings, attentional mechanisms and convolutional neural networks are commonly employed for ABSA tasks. However, because such models lack ways to account for important syntactic restrictions and remote word dependencies, they may identify grammar-independent contextual words as cues for determining aspectual sentiment wrongly. To solve this issue, Zhang [16] suggested an aspect-specific sentiment classification framework (ASGCN) built on the dependency tree of sentences employing syntactic information and word dependencies. Hou [17] stated that GCN models based only on dependency trees are prone to parsing mistakes, thus they combined a self-attentive sequence model with the GCN model to create a novel GCN model based on selective attention (SA-GCN). According to Xiao [18], most techniques cannot adequately capture context semantic information and lack tools to explain important syntactic restrictions and distant lexical dependencies. An attentional coding-based graph convolutional network model (AEGCN) was proposed to address these issues. Zhang and Sun [21] proposed employing graph convolutional networks (GCN) to learn node representations from dependency trees and utilize them in conjunction with other characteristics to classify sentiment. Huang and Carley [22] employed graph attention networks (GAT) to intentionally build word dependencies for similar goals.

The aforementioned GAT-based model does not, however, extract the semantic features and does not take into account the specific aspects and the sentiment relationships between the aspects while creating the adjacency matrix. We suggest a strategy that combines aspect and semantic features with edge feature GAT based on the benefits of GAT in the ABSA challenge. To improve GAT's capacity to precisely capture semantic relationships, we introduce an edge feature.

### 3. Related Methods

Figure 2 illustrates the architecture of the proposed AS-EGAT model, which primarily consists of two elements: (1) The edge feature graph attention network based on aspect syntactic analysis seeks to reinforce the syntactic graph through the context of particular aspect words and the interrelation of different aspects in the context to extract the sentiment features of aspects; (2) The edge feature graph attention network based on semantic features seeks to capture the semantic relationship between words in sentences through self-attention. To produce the sentiment properties of a specific aspect, the feature representations captured by the two components are then merged. We begin by supposing that a sentence has  $n$  words, two aspects, i.e.,

$$s = \{w_1, w_2, \dots, a_{11}, a_{12}, \dots, a_{1p}, \dots, a_{21}, a_{22}, \dots, a_{2p}, \dots, w_n\} \quad (1)$$

where  $w_i$  represents the  $i$ -th contextual word, and  $a_{ij}$  denotes the  $j$ -th word of aspect  $i$ . Every instance has a sentence, one or more aspects, and one or more words for each aspect.

First, we use the pre-trained BERT to initialize the word vector; each word embedding is a distributed representation of a word in the sentence; for a sentence with  $n$  words, we can obtain the corresponding embedding matrix  $X = \{x_1, x_2, \dots, x_n\}$  where  $x_i \in R_m$  is the word embedding of  $w_i$  and  $m$  is the dimension of the word vector.

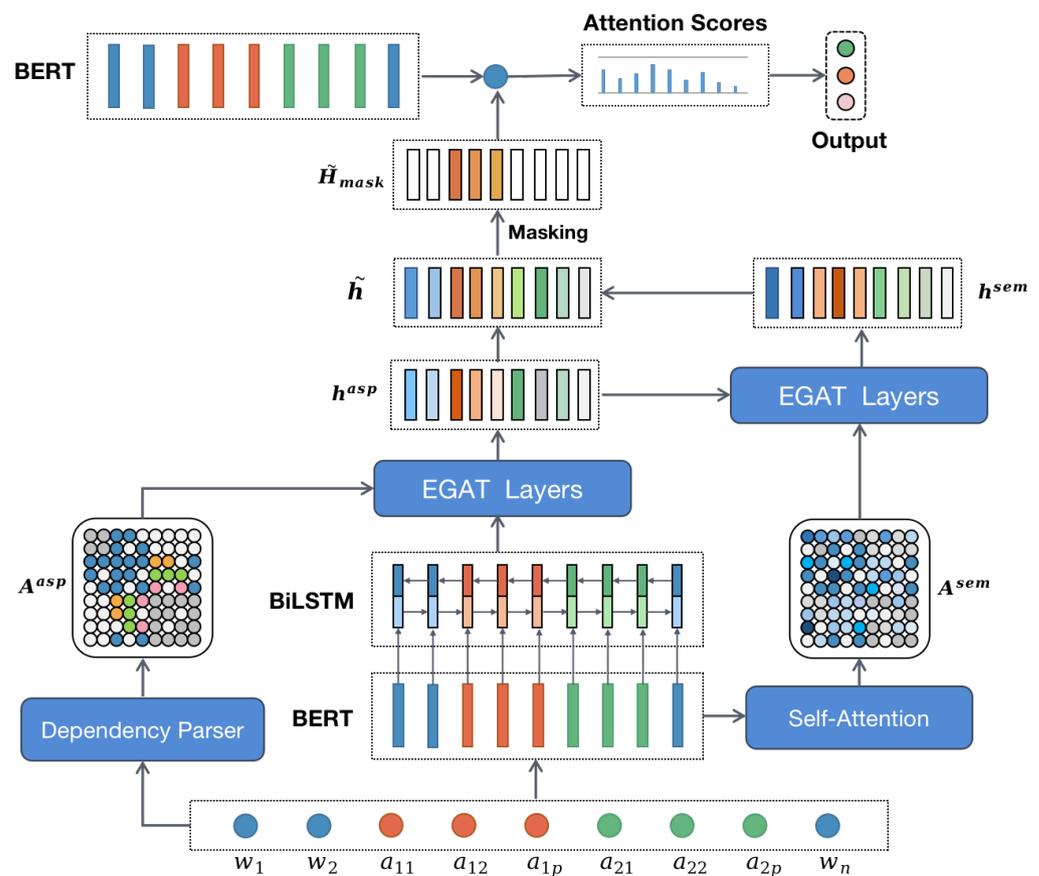


Figure 2. The architecture of the proposed AS-EGAT model.

### 3.1. Aspect Syntactic Graph

To emphasize specific aspects from the context and capture aspect-focused weight maps for that sentence, we calculate the relative position weights of each word and construct a weight map depending on the specific aspect.

$$W_{i,j} = \begin{cases} 1 & \text{if } W_i \in \{a_i^s\} \text{ and } W_j \in \{a_i^s\} \\ \frac{1}{(|j - p^s| + 1)} & \text{if } W_i \in \{a_i^s\} \\ \frac{1}{(|i - p^s| + 1)} & \text{if } W_j \in \{a_i^s\} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

We merge aspect-focused aspect graphs with syntactic analysis dependency trees to produce aspect-focused syntactic dependency adjacency matrices to improve the syntactic reliance of context words and gain relationships between aspects and context words.

$$G_{i,j} = \begin{cases} 1 + W_{i,j} & \text{if } i = j \text{ or } W_i, W_j \text{ in the dependency tree of the sentence} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

Affective expressions in the sentences may not be readily apparent in some aspects, however. In other words, because these aspects' sentiment dependencies are in some ways coupled to other aspects, aspect-focused syntactic analysis graphs may not be sufficient to discover accurate sentiment connections. To make use of the connections between different

aspects in a sentence, we therefore enhance the aspect syntactic graph by including the relative graphs of other aspects in the aspect-focused adjacency matrix.

$$A^{asp} = G + \frac{1}{l} \sum_{a \in \{a_i^o\}} \alpha * G_a \tag{4}$$

$$\alpha = \frac{1}{(|p^o - p^s| + 1)} \tag{5}$$

where  $p^o$  of each  $a \in \{a_i^o\}$  stands for the initial position of another aspect, and  $a_i^o$  is a word set of other aspects of length  $l$ . We create the directionless adjacency matrix, or  $A_{i,j}^{asp} = A_{j,i}^{asp}$ , to enhance the dependency information of the input sentences.

### 3.2. Semantic Graph

The self-attentive process produces an attention matrix as an adjacency matrix. On the one hand, self-attention, which is more flexible than the synchronous strategy structure, can capture semantically relevant sentences for each word in the sentence. The semantic graph, on the other hand, may be applied to the sentiment that is insensitive to syntactic information.

The attention score of each pair of elements was determined in parallel by Self-Attention [23]. The self-attention layer in our proposed AS-EGAT model is used to compute the attention score matrix  $A^{sem} \in \mathbf{R}^{n \times n}$ . Then we take the attention scoring matrix  $A^{sem}$  as the semantic module’s adjacency matrix, namely  $A_{i,j}^{sem} = A_{j,i}^{sem}$ , which can be stated as:

$$A^{sem} = softmax\left(\frac{QW^Q \times (KW^K)^T}{\sqrt{d}}\right) \tag{6}$$

The weight matrices  $W^Q$  and  $W^K$  can be learned, whereas the matrices  $Q$  and  $K$  originate from the output of the BERT layer. The input node feature’s dimension is  $d$ , as well. To determine a sentence’s attention score matrix in our model, we solely employ a self-attention head.

### 3.3. Edge-Featured Graph Attention Networks

We improve the GAT model’s original attention mechanism by making the edge features a significant factor in the calculation of the weight coefficients by the attention mechanism to efficiently mine the edge features of the GAT model, which uses aspect syntactic graphs and semantic graphs as adjacency matrices.

$$\vec{h}_i' = \sigma\left(\frac{1}{k} \sum_{k=1}^k \sum_{j \in N_i} \alpha_{i,j}^k W^k \vec{h}_j\right) \tag{7}$$

$$\alpha_{i,j} = \frac{\exp\left(LeakyReLU\left(\vec{a}^T \left[ W\vec{h}_i \| W\vec{h}_j \| W\vec{A}_{i,j} \right] \right)\right)}{\sum_{k \in N_i} \exp\left(LeakyReLU\left(\vec{a}^T \left[ W\vec{h}_i \| W\vec{h}_j \| W\vec{A}_{i,j} \right] \right)\right)} \tag{8}$$

where  $\vec{a}^T \in \mathbf{R}^{2F'}$  is the feedforward neural network parameter,  $LeakyReLU(\cdot)$  is the activation function of the feedforward neural network,  $\|$  is the concatenation of two features,  $\vec{A}_{i,j}$  is the matching adjacency matrix, and this formula may compute the correlation degree with the neighboring nodes. The number of heads in the attention span of multiple heads is represented by the parameter  $K$ , adding and averaging yields the updated characteristic  $\vec{h}_i'$ .

### 3.4. Training

The initial node of the aspect parsing EGAT layer is formed from the Bi-LSTM layer’s hidden representation, which embeds words as input:

$$H = \{h_1, h_2, \dots, h_n\} = Bi - LSTM(x) \tag{9}$$

The aspect parsing EGAT layer’s final representation is  $h^{asp}$ .  $h^{asp}$  provides the initial node for the EGAT layer based on the semantic feature, while  $h^{sem}$  provides the final representation for the EGAT layer based on the semantic feature. To extract the relationship between aspect syntactic and semantic information, we merge these two final representations.

$$\tilde{h} = h^{asp} + \lambda h^{sem} \tag{10}$$

where  $\lambda$  is the coefficient of semantic features. We employ aspect-specific masking to cover non-aspect representations, highlighting the important aspects of aspect words:

$$\tilde{H}_{mask} = \{0, \dots, \tilde{h}_\tau, \dots, \tilde{h}_{\tau+k-1}, \dots, 0\} \tag{11}$$

where  $\tilde{h}_t$  is the representation of the  $t$ -th word that AS-EGAT learned,  $\tau$  is the index at which a certain aspect first appears, and  $k$  is the length of that aspect. We then use a retrieving-based attention method, inspired by Zhang [16], to extract significant sentiment features from contextual representations of particular aspects:

$$\xi_t = \sum_{i=1}^n h_i^c T \tilde{h}_i = \sum_{i=\tau}^{\tau+k-1} h_i^c T \tilde{h}_i \tag{12}$$

$$\eta_t = \frac{\exp(\xi_t)}{\sum_{i=1}^n \exp(\xi_i)} \tag{13}$$

As a result, the following is the final representation of the input for specific aspects:

$$y = softmax(W_o r + b_o) \tag{14}$$

$$r = \sum_{i=1}^n \eta_i h_i^c \tag{15}$$

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

Minimizing the cross entropy loss between the predicted distribution and the ground real distribution is the definition of the classifier training objective:

$$L = - \sum_{i=1}^S \sum_{j=1}^C \hat{y}_i^j \cdot \log(y_i^j) + \gamma \|\Theta\|_2 \tag{16}$$

where  $C$  is the classification number and  $S$  is the training scale. The actual sentiment distribution is  $\hat{y}$ . The regularization term  $L_2$  weight is  $\gamma$ . All trainable parameters are represented by  $\Theta$ .

## 4. Experiment

### 4.1. Dataset and Experimental Setup

In the ABSA, since the LAP14, REST15, REST16, and MAMS datasets are the most commonly used and popular benchmark datasets and have abundant sources, we chose them to test the universality of our model. The LAP14, REST15, and REST16 datasets are the laptop and restaurant review datasets commonly used in the SemEval 2014 (Laptop 14) [24], SemEval 2015 (Restaurant 15) [25], and SemEval 2016 (Restaurant 16) [26] tasks. The MAMS

datasets [27] are multiple aspects. To further demonstrate the validity of our proposed model, we use the dataset provided by Mukherjee [28], which uses similar guidelines to label two new datasets from the field of e-commerce, namely T-shirts and Television [28]. Each sample consists of a comment sentence, an aspect (single or multiple words), and a corresponding aspect of emotional polarity (positive or neutral, or negative). The statistics information of these six datasets is shown in Table 1.

In our experiment, each word was initialized into a 768-dimensional word vector using the BERT preprocessing model, and the  $\gamma$  of the  $L_2$  regularization coefficient was set to  $10^{-5}$ . With a learning rate of  $10^{-3}$ , Adam was employed as an optimizer to train the model. The scale of batch training was 16. We use a uniform distribution to initialize all of the  $W$  and  $b$  at random.

**Table 1.** The statistical information of all the datasets.

Dataset	Positive		Neutral		Negative	
	Train	Test	Train	Test	Train	Test
LAP14	994	341	464	169	870	128
REST15	1178	439	50	35	382	328
REST16	1620	597	88	38	709	190
MAMS	3380	400	5042	607	2764	329
T-shirt	1122	270	50	16	699	186
Television	2540	618	287	67	919	257

#### 4.2. Comparative Model

We compared the proposed model (AS-EGAT) with the following models:

- **TD-LSTM** [5]: develops two target-dependent long short-term memory (LSTM) models, where target information is automatically taken into account.
- **ATAE-LSTM** [11]: proposes an Attention-based Long Short-Term Memory Network for aspect-level sentiment classification. The attention mechanism can concentrate on different parts of a sentence when different aspects are taken as input.
- **MemNet** [29]: introduces a deep memory network for aspect level sentiment classification. This approach explicitly captures the importance of each context word.
- **IAN** [15]: proposes the interactive attention networks (IAN) to interactively learn attentions in the contexts and targets, and generate the representations for targets and contexts separately.
- **RAM** [30]: adopts a multiple-attention mechanism to capture sentiment features separated by a long distance, so that it is more robust against irrelevant information.
- **GCAE** [3]: proposes a model based on convolutional neural networks and gating mechanisms, which is more accurate and efficient.
- **MGAN** [31]: proposes a fine-grained attention mechanism, which can capture the word-level interaction between aspect and context. It then leverages the fine-grained and coarse-grained attention mechanisms to compose the MGAN framework.
- **AOA** [14]: models aspects and sentences jointly and explicitly captures the interaction between aspects and context sentences.
- **TNet-LF** [2]: proposes a component to generate target-specific representations of words in the sentence, meanwhile incorporating a mechanism for preserving the original contextual information from the RNN layer.
- **ASGCN-DT** [16] and **ASGCN-DG** [16]: propose to build a Graph Convolutional Network (GCN) over the dependency tree of a sentence to exploit syntactical information and word dependencies.
- **BERT** [8]: designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers.

- **CapsNet+BERT** [27]: presents a new large-scale Multi-Aspect Multi-Sentiment (MAMS) dataset, in which each sentence contains at least two different aspects with different sentiment polarities.
- **GIN+BERT** [32]: adopts two attention-based networks to learn the contextual sentiment for the entity and attribute independently and interactively. Further, based on the interactive attentions learned from entities and attributes, the coordinative gate units are exploited to reconcile and purify the sentiment features for the aspect sentiment prediction.
- **MIMLLN+BERT** [33]: proposes a Multi-Instance Multi-Label Learning Network for Aspect-Category sentiment analysis (AC-MIMLLN), which treats sentences as bags, words as instances, and the words indicating an aspect category as the key instances of the aspect category.
- **SGGCN+BERT** [34]: proposes a mechanism to obtain the importance scores for each word in the sentences based on the dependency trees that are then injected into the model to improve the representation vectors for ABSA.
- **DGEDT+BERT** [35]: proposes a dependency graph enhanced dual-transformer network (named DGEDT) by jointly considering the flat representations learned from Transformer and graph-based representations learned from the corresponding dependency graph in an iterative interaction manner.
- **R-GAT+BERT** [36]: defines a unified aspect-oriented dependency tree structure rooted at a target aspect by reshaping and pruning an ordinary dependency parse tree.
- **DeBERTa** [37]: presents a new pre-trained language model, DeBERTaV3, which improves the original DeBERTa model by replacing mask language modeling (MLM) with replaced token detection (RTD), a more sample-efficient pre-training task.
- **T-GCN+BERT** [38]: proposes an approach to explicitly utilize dependency types for ABSA with type-aware graph convolutional networks (T-GCN), where attention is used in T-GCN to distinguish different edges (relations) in the graph, and attentive layer ensemble is proposed to comprehensively learn from different layers of T-GCN.
- **DualGCN+BERT** [39]: proposes a dual graph convolutional networks (DualGCN) model that considers the complementarity of syntax structures and semantic correlations simultaneously.
- **GL-GCN** [40]: proposes a novel aspect-based sentiment classification approach, i.e., Global and Local Dependency Guided Graph Convolutional Networks (GL-GCN).
- **SenticGCN+BERT** [41]: proposes a graph convolutional network based on SenticNet to leverage the effective dependencies of the sentence according to the specific aspect, called Sentic GCN.
- **SEDC-GCN** [42]: proposes a novel GCN based model, named the Structure-Enhanced Dual-Channel Graph Convolutional Network (SEDC-GCN).
- **HGCN** [43]: proposes a hybrid graph convolutional network (HGCN) to synthesize information from constituency tree and dependency tree, exploring the potential of linking two syntax parsing methods to enrich the representation.
- **AEN+BERT** [44]: proposes an Attentional Encoder Network (AEN) which eschews recurrence and employs attention based encoders for the modeling between context and target.
- **STGNN-GRU** [45]: proposes a graph Fourier transform based network with features created in the spectral domain. Fourier transform is used to switch to the frequency (spectral) domain where new features are created.
- **LGCF-CDM** [46] and **LGCF-CDW** [46]: propose a multilingual learning model based on the interactive learning of local and global context focus, namely LGCF. This model can effectively learn the correlation between local context and target aspects and the correlation between global context and target aspects simultaneously.

### 4.3. Evaluation Metric

In order to evaluate the performance of the model, Accuracy (Acc.) and Macro-F1 score (F1) is selected as the main evaluation metric. Accuracy can intuitively see the percentage of predicted correct results in the total sample. However, for the data with an unbalanced proportion of positive and negative samples, the category with a large proportion will affect the accuracy. The F1 score is the harmonic average of Recall and Precision, which is a comprehensive evaluation metric.

### 4.4. Experimental Results

Table 2 shows an overview of the experimental results for LAP14, REST15, REST16, and MAMS datasets using Acc. and F1 metrics. The experimental results show that our proposed AS-EGAT model is superior to the current state-of-the-art baselines. More specifically, Acc./F1 scores in our AS-EGAT model improved to 0.76%/1.16%, 0.29%/1.16%, 0.05%/1.23%, and 0.15%/0.37% compared to baseline models for LAP14, REST15, REST16, and MAMS.

Table 3 shows an overview of the results of experiments on the T-shirt and Television datasets using Acc. and F1 metrics. Acc./F1 scores improved by 0.22%/0.53% and 0.38%/1.93% in our AS-EGAT model compared with baseline models for T-shirt and Television.

This verifies the validity of AS-EGAT in ABSA. It is also worth noting that the performance of AS-EGAT is significantly better than that of previous models based on the graph convolution network and graph attention network, which fundamentally verifies the validity of the new solution of graph construction used in this work.

**Table 2.** Performance comparison on different models on LAP14, REST15, REST16, and MAMS datasets. The best performance is bold-typed.

Model	LAP14		REST15		REST16		MAMS	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
TD-LSTM (2016a) [5]	71.83	68.43	76.39	58.70	82.16	54.21	74.59	–
ATAE-LSTM (2016b) [11]	68.88	63.93	78.48	60.53	83.77	61.71	77.05	–
MenNet (2016b) [29]	70.64	65.17	77.31	58.28	85.44	65.99	64.56	–
IAN (2017) [15]	72.05	67.38	78.54	52.65	84.74	55.21	76.60	–
RAM (2017) [30]	74.49	71.35	79.98	60.57	83.88	62.14	–	–
GCAE (2018) [3]	71.98	68.71	77.56	56.03	83.70	62.69	77.59	–
MGAN (2018) [31]	75.39	72.47	79.36	57.26	87.06	62.29	–	–
AOA (2018) [14]	72.62	67.52	78.17	57.02	87.50	66.21	77.26	–
TNet-LF (2018) [2]	74.61	70.14	78.47	59.47	89.07	70.43	–	–
ASGCN-DT (2019) [16]	74.14	69.24	79.34	60.78	88.69	66.64	–	–
ASGCN-DG (2019) [16]	75.55	71.05	79.89	61.89	88.99	67.48	–	–
BERT (2019) [8]	77.59	73.28	83.48	66.18	90.10	74.16	80.62	80.77
CapsNet+BERT (2019) [27]	–	–	81.89	61.85	86.50	62.12	82.97	–
GIN+BERT (2020) [32]	–	–	83.96	66.03	89.47	74.87	–	–
MIMLLN+BERT (2020) [33]	–	–	82.76	65.10	88.12	73.05	–	–
SGGCN+BERT (2020) [34]	–	–	82.72	65.86	90.52	74.53	–	–
DGEDT+BERT (2020) [35]	<b>79.80</b>	<b>75.60</b>	79.89	61.89	88.99	67.48	–	–
R-GAT+BERT (2020) [36]	79.73	75.50	–	–	91.87	75.54	81.75	80.87
DeBERTa (2021) [37]	–	–	–	–	–	–	83.06	82.52
T-GCN+BERT (2021) [38]	–	–	85.26	<b>71.69</b>	<b>92.32</b>	77.29	<b>83.68</b>	<b>83.07</b>
DualGCN+BERT (2021) [39]	79.51	–	83.78	–	91.43	–	–	–
GL-GCN (2021) [40]	76.91	72.76	80.81	64.99	88.47	69.64	–	–
SenticGCN+BERT (2022) [41]	78.21	74.07	<b>85.32</b>	71.28	91.97	<b>79.56</b>	–	–
SEDC-GCN (2022) [42]	77.74	74.68	81.73	66.23	90.75	73.84	–	–
HGCN (2022) [43]	78.82	–	80.81	–	88.92	–	–	–
AS-EGAT+BERT (ours)	<b>80.56</b>	<b>76.76</b>	<b>85.61</b>	<b>72.85</b>	<b>92.37</b>	<b>80.79</b>	<b>83.83</b>	<b>83.44</b>

**Table 3.** Performance comparison on different models on T-shirt and Television datasets. The best performance is bold-typed.

Model	T-Shirt		Television	
	Acc.	F1	Acc.	F1
BERT-SPC (2019) [8]	93.13	73.86	<b>89.96</b>	<b>74.68</b>
BERT-AEN (2019) [8]	88.69	72.25	87.09	67.92
ASGCN-DT (2019) [16]	92.01	71.86	84.96	69.15
ASGCN-DG (2019) [16]	91.59	73.09	85.27	69.40
ASGCN+BERT (2019) [16]	92.22	77.24	88.99	74.42
AEN+BERT (2019) [44]	88.69	72.25	87.09	67.92
STGNN-GRU (2022) [45]	–	–	89.73	75.73
LGCF-CDM (2022) [46]	93.22	76.73	–	–
LGCF-CDW (2022) [46]	<b>93.64</b>	<b>77.23</b>	–	–
AS-EGAT+BERT (ours)	<b>93.86</b>	<b>77.76</b>	<b>90.34</b>	<b>76.61</b>

#### 4.5. Ablation Study

To confirm the viability of our suggested strategy, we also carried out ablation tests to examine the effects of different modules in the AS-EGAT model. Table 4 presents the outcomes. It can be shown that (1) eliminating the “aspect syntactic features” module lowers the model’s performance, proving that adding the “aspect syntactic features” information to the model improves aspect-oriented sentiment analysis; (2) the module’s unsatisfactory performance on all datasets when “semantic characteristics” are removed is evidence that rich grammatical knowledge can decrease dependency parsing errors; (3) the “edge feature” module can improve the correlation of edges with a strong correlation, as shown by the fact that removing it causes the model’s performance to decline; (4) the model effect becomes worse when the “aspect syntactic features and side features” module is removed, demonstrating the module’s influence on the model; (5) the performance of the model becomes unsatisfactory when the “semantic feature and edge feature” module is removed, demonstrating the module’s significance to the model. Our AS-EGAT model delivers the finest performance overall.

**Table 4.** Experimental results of the ablation study.

Model	LAP14		REST15		REST16		MAMS	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
AS-EGAT+BERT w/o A	80.25	76.02	85.24	71.02	91.56	80.41	83.01	82.38
AS-EGAT+BERT w/o S	79.47	76.15	84.87	69.05	91.56	73.52	83.46	83.05
AS-EGAT+BERT w/o E	78.06	73.83	85.16	71.04	91.23	74.77	83.53	82.92
AS-EGAT+BERT w/o A+E	79.15	75.01	85.06	65.81	90.91	76.73	83.68	83.04
AS-EGAT+BERT w/o S+E	79.47	76.42	85.06	71.15	90.91	73.38	83.76	83.39
AS-EGAT+BERT	<b>80.56</b>	<b>76.76</b>	<b>85.61</b>	<b>72.85</b>	<b>92.37</b>	<b>80.79</b>	<b>83.83</b>	<b>83.44</b>

#### 4.6. Influence of the Coefficient of Semantic

We conducted trials with various  $\lambda$  values to further examine the impact of the semantic module in the AS-EGAT model, and the outcomes are displayed in Figure 3. It can be seen that the datasets perform best when the  $\lambda$  value reaches 1.0. The curve swings significantly whether  $\lambda$  is less than or more than 1.0, indicating that the best semantic information may be recovered when  $\lambda$  is equal to 1.0.

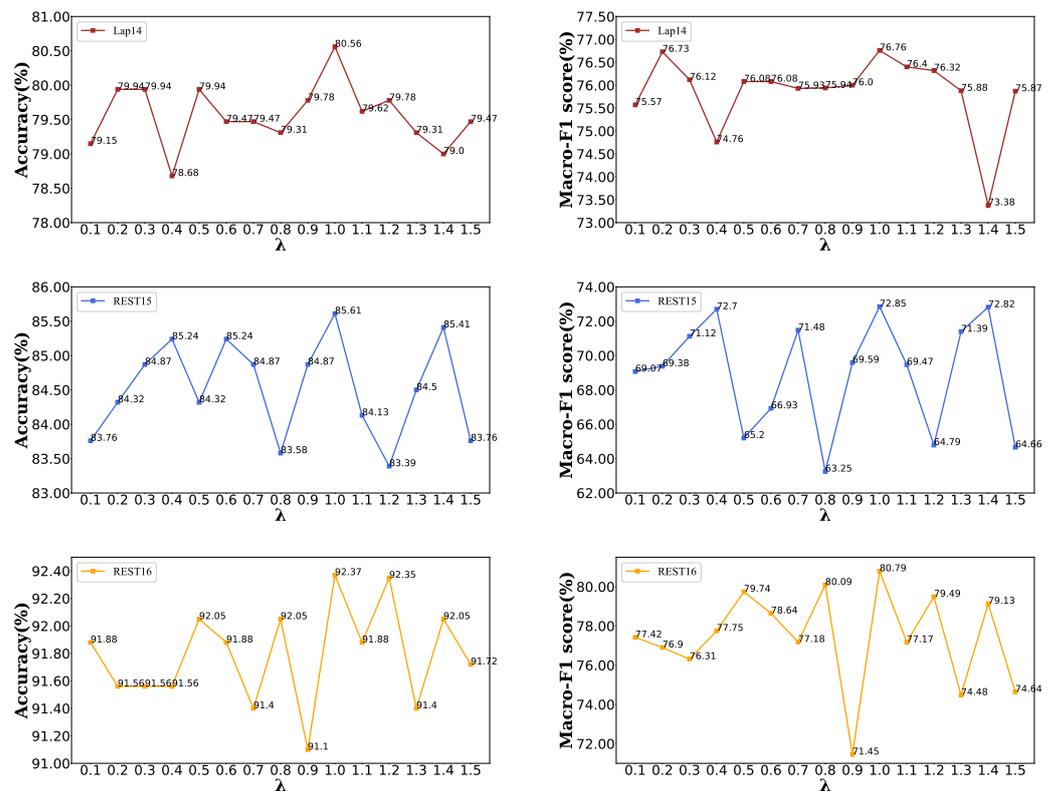


Figure 3. Impact of semantic coefficients. Accuracy and Macro-F1 score based on different values of  $\lambda$  is reported.

#### 4.7. Analysis of Different Sentiment Polarity

Figure 4 shows the performance of the AS-EGAT model in different sentimental polarities. In the five datasets of LAP14, REST15, REST16, T-shirt, and Television, “positive” has the best performance, while “neutral” has the worst performance. The results were similar for the three sentimental polarities in the “MAMS” dataset. This phenomenon is related to the distribution of sentimental polarity in the dataset. The number of “neutral” training data in the “MAMS” dataset is higher than the other two sentimental polarities, while for the other five datasets, the number of “neutral” training data is the least. Furthermore, the model struggles to capture emotional information in “neutral” sentences, resulting in low performance.

#### 4.8. Single and Multiple Aspects Analysis

In order to further analyze the effects of the AS-EGAT model in single and multiple aspects, we divided the test set into single and multiple aspects according to the number of aspects in the sentence for experiments, and the results are shown in Figure 5. According to the experimental results, it can be found that in LAP14, REST15, and MAMS datasets, single and multiple aspects accuracy rates are basically the same, while in REST16, T-shirt, and Television datasets, multiple aspects accuracy rates are slightly higher than single aspects accuracy rates. From this, we can see that the AS-EGAT model is more effective in multiple aspects.

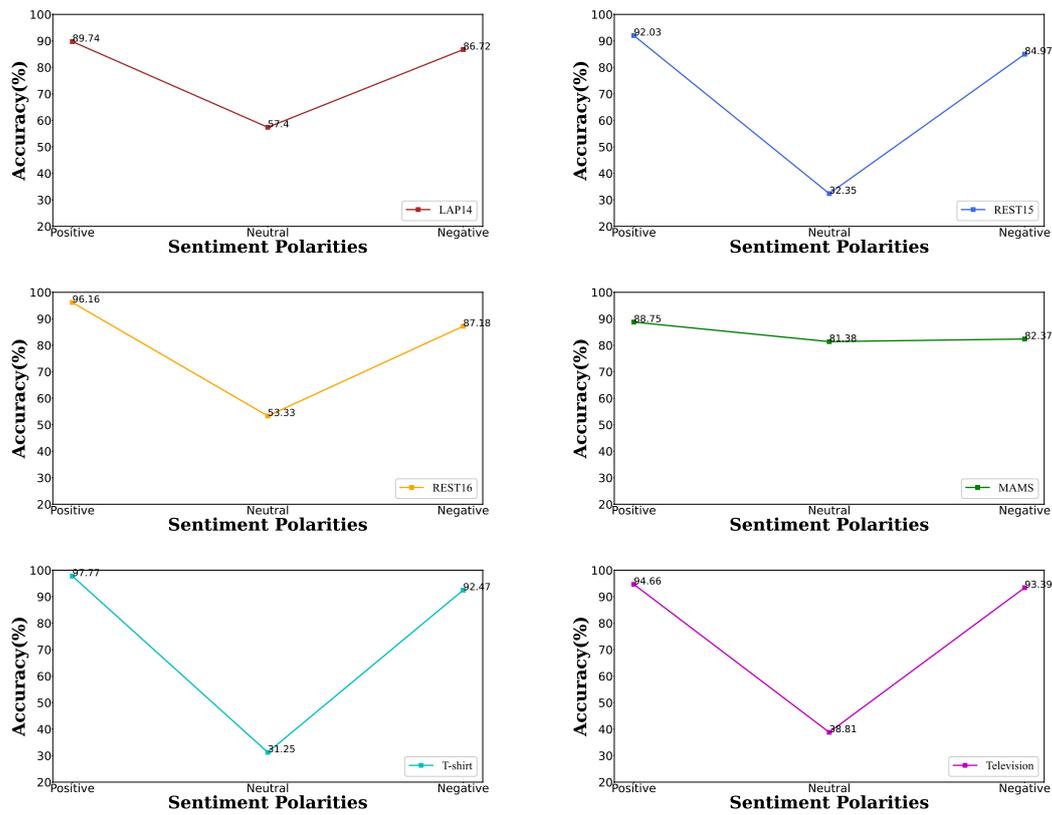


Figure 4. Experimental results of different sentiment polarity.

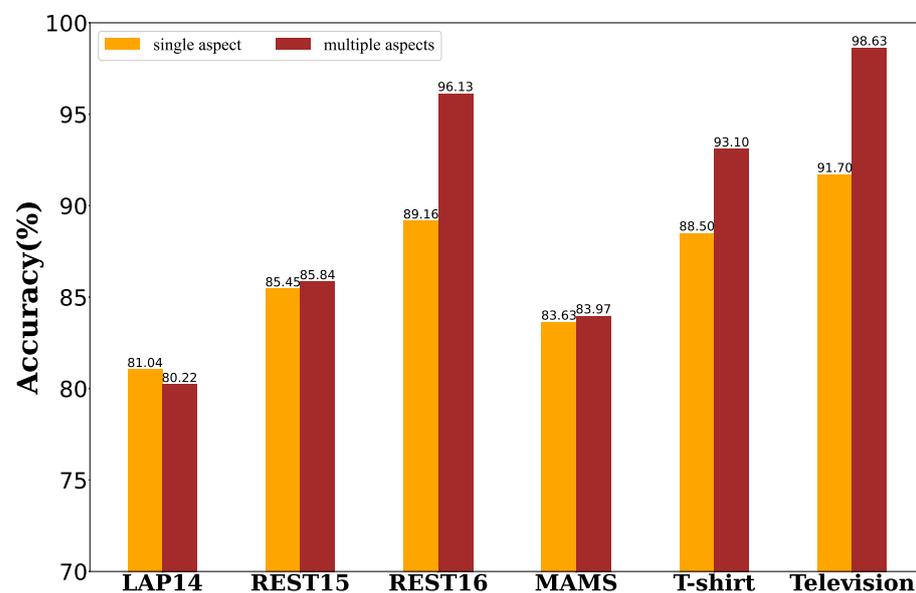
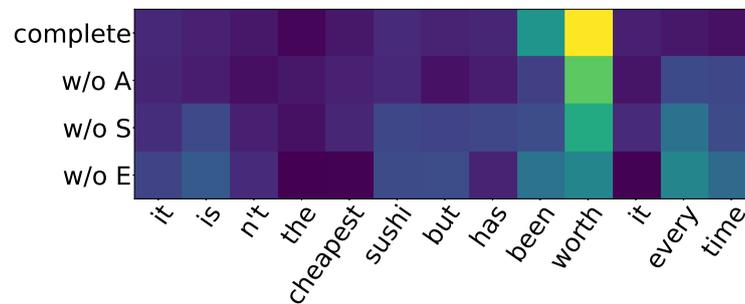


Figure 5. Experimental results of single and multiple aspects.

4.9. Visualization

We exhibit two visual attention weight graphs in Figure 6 to demonstrate how the AS-EGAT model improves the performance of aspect sentiment analysis. As shown in Figure 6a, the weight value of the relevant terms in the modules of ‘aspect syntactic analysis’, ‘semantic analysis’, and ‘edge feature’ is less than that of AS-EGAT. As a result, the AS-EGAT proposed can pay more attention to important aspect aspects in order to extract the sentiment characteristics of specific aspects. The multi-faceted examples in Figure 6b

show how different aspects of the same sentence receive varied weights of attention, demonstrating the efficiency of our model in extracting sentiment features from diverse aspects of the text.



(a) Example of ablation study attention



(b) Examples of multiple aspects of attention

**Figure 6.** The attention visualizations.

## 5. Conclusions

In this paper, we explore a new ABSA solution for constructing aspect syntactic analysis graphs and semantic dependency graphs. On this basis, an attention network (AS-EGAT) model based on aspect syntactic and semantic features is proposed to extract aspect-based sentiment features from the aspect and semantic perspectives. Therefore, the proposed AS-EGAT model can pay significant attention to keywords when processing multiple words, connect valuable sentiment features of related aspects when considering aspects without obvious sentiment expression, and obtain corresponding sentiment features by extracting semantic features for sentences insensitive to the syntactic dependency tree. Experimental results on six datasets show that the proposed AS-EGAT can outperform existing graph-based models and BERT-based models.

Notably, the main contribution of this paper is to increase sentence structure information by constructing syntactic analysis graphs and semantic dependency graphs, a general strategy that can be easily applied to these BERT and graph-based neural network methods. From the experimental data, we can find that the model has poor performance for sentences with “neutral” sentiment polarity. For this purpose, in future work, we will construct new graphs to capture the sentiment information of sentences with “neutral” sentiment polarity and continue investigating its validity in these BERT-based frameworks.

**Author Contributions:** conceptualization, L.S. and J.L.; methodology, L.S.; writing—original draft, L.S.; writing—review and editing, J.L. and G.L.; supervision, J.L., L.S., G.L. and X.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the 2022 Central Government Guided Local Development Science and Technology Special Project (2022L3029).

**Data Availability Statement:** The LAP14 dataset is available at <https://alt.qcri.org/semEval2014/task4/index.php>, accessed on 6 May 2022; the REST15 dataset is available at <https://alt.qcri.org/semEval2015/task12/index.php>, accessed on 6 May 2022; REST16 dataset is available at <https://alt.qcri.org/semEval2016/task5/index.php>, accessed on 6 May 2022 and the MAMS, T-shirt and Television datasets are available at <https://github.com/yangheng95/ABSADatasets/tree/v2>

[.0/datasets/apc\\_datasets](#), accessed on 16 January 2023. Experimental data and code related to this paper can be obtained by contacting the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fan, C.; Gao, Q.; Du, J.; Gui, L.; Xu, R.; Wong, K.F. Convolution-Based Memory Network for Aspect-Based Sentiment Analysis. In *SIGIR '18, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, Ann Arbor, MI, USA, 8–12 July 2018*; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1161–1164. [[CrossRef](#)]
2. Li, X.; Bing, L.; Lam, W.; Shi, B. Transformation Networks for Target-Oriented Sentiment Classification. *arXiv* **2018**, arXiv:1805.01086.
3. Xue, W.; Li, T. Aspect Based Sentiment Analysis with Gated Convolutional Networks. *arXiv* **2018**, arXiv:1805.07043.
4. Huang, B.; Carley, K.M. Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification. *arXiv* **2019**, arXiv:1909.06276.
5. Tang, D.; Qin, B.; Feng, X.; Liu, T. Target-Dependent Sentiment Classification with Long Short Term Memory. *arXiv* **2015**, arXiv:1512.01100.
6. Zhang, M.; Zhang, Y.; Vo, D. Gated Neural Networks for Targeted Sentiment Analysis. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AR, USA, 12–17 February 2016*; Schuurmans, D., Wellman, M.P., Eds.; AAAI Press: Palo Alto, CA, USA, 2016; pp. 3087–3093.
7. Ruder, S.; Ghaffari, P.; Breslin, J.G. A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis. *arXiv* **2016**, arXiv:1609.02745.
8. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv* **2018**, arXiv:1810.04805.
9. Xu, H.; Liu, B.; Shu, L.; Yu, P.S. BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis. *arXiv* **2019**, arXiv:1904.02232.
10. Sun, C.; Huang, L.; Qiu, X. Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. *arXiv* **2019**, arXiv:1903.09588.
11. Wang, Y.; Huang, M.; Zhu, X.; Zhao, L. Attention-based LSTM for Aspect-level Sentiment Classification. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, TX, USA, 1–5 November 2016*; Association for Computational Linguistics: Austin, TX, USA, 2016; pp. 606–615. [[CrossRef](#)]
12. Yang, M.; Tu, W.; Wang, J.; Xu, F.; Chen, X. Attention-Based LSTM for Target-Dependent Sentiment Classification. In *AAAI'17, Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, VA, USA, 4–9 February 2017*; AAAI Press: Palo Alto, CA, USA, 2017; pp. 5013–5014.
13. Liu, J.; Zhang, Y. Attention Modeling for Targeted Sentiment. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, Valencia, Spain, 3–7 April 2017*; Association for Computational Linguistics: Valencia, Spain, 2017; pp. 572–577.
14. Huang, B.; Ou, Y.; Carley, K.M. Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks. *arXiv* **2018**, arXiv:1804.06536.
15. Ma, D.; Li, S.; Zhang, X.; Wang, H. Interactive Attention Networks for Aspect-Level Sentiment Classification. *arXiv* **2017**, arXiv:1709.00893.
16. Zhang, C.; Li, Q.; Song, D. Aspect-based Sentiment Classification with Aspect-specific Graph Convolutional Networks. *arXiv* **2019**, arXiv:1909.03477.
17. Hou, X.; Huang, J.; Wang, G.; Huang, K.; He, X.; Zhou, B. Selective Attention Based Graph Convolutional Networks for Aspect-Level Sentiment Classification. *arXiv* **2019**, arXiv:1910.10857.
18. Xiao, L.; Hu, X.; Chen, Y.; Xue, Y.; Gu, D.; Chen, B.; Zhang, T. Targeted Sentiment Classification Based on Attentional Encoding and Graph Convolutional Networks. *Appl. Sci.* **2020**, *10*, 957. [[CrossRef](#)]
19. Kipf, T.N.; Welling, M. Semi-Supervised Classification with Graph Convolutional Networks. *arXiv* **2016**, arXiv:1609.02907.
20. Veličković, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph Attention Networks. *arXiv* **2017**, arXiv:1710.10903.
21. Sun, K.; Zhang, R.; Mensah, S.; Mao, Y.; Liu, X. Aspect-Level Sentiment Analysis Via Convolution over Dependency Tree. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019*; Association for Computational Linguistics: Hong Kong, China, 2019; pp. 5679–5688. [[CrossRef](#)]
22. Huang, B.; Carley, K.M. Syntax-Aware Aspect Level Sentiment Classification with Graph Attention Networks. *arXiv* **2019**, arXiv:1909.02606.
23. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.U.; Polosukhin, I. Attention is All you Need. In *Advances in Neural Information Processing Systems*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: New York, NY, USA, 2017; Volume 30.
24. Pontiki, M.; Galanis, D.; Pavlopoulos, J.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, 23–24 August 2014*; Association for Computational Linguistics: Dublin, Ireland, 2014; pp. 27–35. [[CrossRef](#)]

25. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Manandhar, S.; Androutsopoulos, I. SemEval-2015 Task 12: Aspect Based Sentiment Analysis. In Proceedings of the SemEval@NAACL-HLT, Denver, CO, USA, 4–5 June 2015; Cer, D.M., Jurgens, D., Nakov, P., Zesch, T., Eds.; The Association for Computer Linguistics: Stroudsburg, PA, USA, 2015; pp. 486–495.
26. Pontiki, M.; Galanis, D.; Papageorgiou, H.; Androutsopoulos, I.; Manandhar, S.; Al-Smadi, M.; Al-Ayyoub, M.; Zhao, Y.; Qin, B.; De Clercq, O.; et al. SemEval-2016 Task 5: Aspect Based Sentiment Analysis. In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016; Association for Computational Linguistics: San Diego, CA, USA, 2016; pp. 19–30. [[CrossRef](#)]
27. Jiang, Q.; Chen, L.; Xu, R.; Ao, X.; Yang, M. 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 2019, Hong Kong, China, 3–7 November 2019; Inui, K., Jiang, J., Ng, V., Wan, X., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2019; pp. 6279–6284. [[CrossRef](#)]
28. Mukherjee, R.; Shetty, S.; Chattopadhyay, S.; Maji, S.; Datta, S.; Goyal, P. Reproducibility, Replicability and Beyond: Assessing Production Readiness of Aspect Based Sentiment Analysis in the Wild. In Proceedings of the Advances in Information Retrieval, Virtual Event, 28 March–1 April 2021; Hiemstra, D., Moens, M.F., Mothe, J., Perego, R., Potthast, M., Sebastiani, F., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 92–106.
29. Tang, D.; Qin, B.; Liu, T. Aspect Level Sentiment Classification with Deep Memory Network. *arXiv* **2016**, arXiv:1605.08900.
30. Chen, P.; Sun, Z.; Bing, L.; Yang, W. Recurrent Attention Network on Memory for Aspect Sentiment Analysis. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Copenhagen, Denmark, 7–11 September 2017; Association for Computational Linguistics: Copenhagen, Denmark, 2017; pp. 452–461. [[CrossRef](#)]
31. Fan, F.; Feng, Y.; Zhao, D. Multi-grained Attention Network for Aspect-Level Sentiment Classification. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 31 October–4 November 2018; Association for Computational Linguistics: Brussels, Belgium, 2018; pp. 3433–3442. [[CrossRef](#)]
32. Yin, R.; Su, H.; Liang, B.; Du, J.; Xu, R. Extracting the Collaboration of Entity and Attribute: Gated Interactive Networks for Aspect Sentiment Analysis. In Proceedings of the Natural Language Processing and Chinese Computing, Zhengzhou, China, 14–18 October 2020; Zhu, X., Zhang, M., Hong, Y., He, R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 802–814.
33. Li, Y.; Yin, C.; Zhong, S.; Pan, X. Multi-Instance Multi-Label Learning Networks for Aspect-Category Sentiment Analysis. *arXiv* **2020**, arXiv:2010.02656.
34. Veyseh, A.P.B.; Nour, N.; Derroncourt, F.; Tran, Q.H.; Dou, D.; Nguyen, T.H. Improving Aspect-based Sentiment Analysis with Gated Graph Convolutional Networks and Syntax-based Regulation. *arXiv* **2020**, arXiv:2010.13389.
35. Tang, H.; Ji, D.; Li, C.; Zhou, Q. Dependency Graph Enhanced Dual-transformer Structure for Aspect-based Sentiment Classification. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020; Association for Computational Linguistics: Stroudsburg, PA, USA, 2020; pp. 6578–6588. [[CrossRef](#)]
36. Wang, K.; Shen, W.; Yang, Y.; Quan, X.; Wang, R. Relational Graph Attention Network for Aspect-based Sentiment Analysis. *arXiv* **2020**, arXiv:2004.12362.
37. He, P.; Gao, J.; Chen, W. DeBERTaV3: Improving DeBERTa using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing. *arXiv* **2021**, arXiv:2111.09543.
38. Tian, Y.; Chen, G.; Song, Y. Aspect-based Sentiment Analysis with Type-aware Graph Convolutional Networks and Layer Ensemble. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Online, 6–11 June 2021; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 2910–2922. [[CrossRef](#)]
39. Li, R.; Chen, H.; Feng, F.; Ma, Z.; Wang, X.; Hovy, E. 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), Online, 1–6 August 2021; Association for Computational Linguistics: Stroudsburg, PA, USA, 2021; pp. 6319–6329. [[CrossRef](#)]
40. Zhu, X.; Zhu, L.; Guo, J.; Liang, S.; Dietze, S. GL-GCN: Global and Local Dependency Guided Graph Convolutional Networks for aspect-based sentiment classification. *Expert Syst. Appl.* **2021**, *186*, 115712. [[CrossRef](#)]
41. Liang, B.; Su, H.; Gui, L.; Cambria, E.; Xu, R. Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowl.-Based Syst.* **2022**, *235*, 107643. [[CrossRef](#)]
42. Zhu, L.; Zhu, X.; Guo, J.; Dietze, S. Exploring rich structure information for aspect-based sentiment classification. *J. Intell. Inf. Syst.* **2022**, 1–21. [[CrossRef](#)]
43. Xu, L.; Pang, X.; Wu, J.; Cai, M.; Peng, J. Learn from structural scope: Improving aspect-level sentiment analysis with hybrid graph convolutional networks. *Neurocomputing* **2023**, *518*, 373–383. [[CrossRef](#)]
44. Song, Y.; Wang, J.; Jiang, T.; Liu, Z.; Rao, Y. Targeted Sentiment Classification with Attentional Encoder Network. In Proceedings of the Artificial Neural Networks and Machine Learning–ICANN 2019: Text and Time Series, Munich, Germany, 17–19 September 2019; Tetko, I.V., Kůrková, V., Karpov, P., Theis, F., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 93–103.
45. Chakraborty, A. Aspect Based Sentiment Analysis Using Spectral Temporal Graph Neural Network. *arXiv* **2022**, arXiv:2202.06776.
46. He, J.; Wumaier, A.; Kadeer, Z.; Sun, W.; Xin, X.; Zheng, L. A Local and Global Context Focus Multilingual Learning Model for Aspect-Based Sentiment Analysis. *IEEE Access* **2022**, *10*, 84135–84146. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.