Jointly Learning Aspect-Focused and Inter-Aspect Relations with Graph Convolutional Networks for Aspect Sentiment Analysis

Bin Liang[1], Rongdi Yin[1], Lin Gui[2*], Jiachen Du[1], Ruifeng Xu[1*]
[1]School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), China
[2]Department of Computer Science, University of Warwick, Coventry, U.K
(* Corresponding Author. Source code: https://github.com/BinLiang-NLP/InterGCN-ABSA)

INTRODUCTION

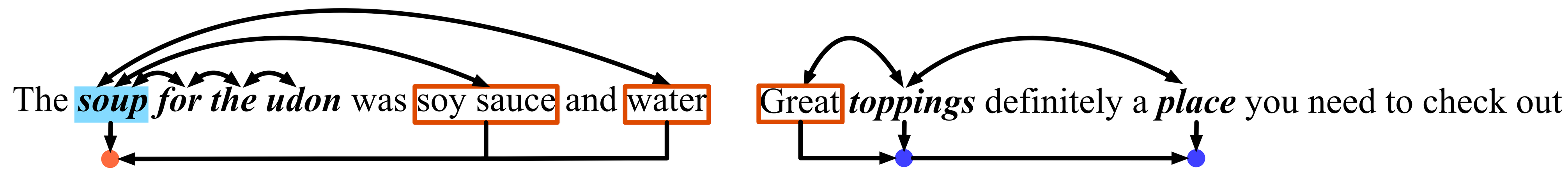
Problem Definition: Aspect sentiment analysis aims to identify the sentiment polarity (e.g. positive, negative, or neutral) towards a given aspect (term) in a sentence. For example, given the aspects: *food* and *service*, and a sentence: *great food but the service is dreadful*, the sentiment polarity of aspect *food* is positive, while for the aspect *service* is negative.

Main Challenge:

- Aspect sentiment analysis aims to discriminate sentiment polarities according to different aspects.
- Aspect may contain no explicit sentiment, in which, sentiment polarities need to be detected according to sentiment relations with other aspects.

MOTIVATION & CONTRIBUTION

Motivation: In Example (a), we could identify the aspect word "*soup*" is the key aspect word of this aspect. In Example (b), the sentiment polarity of aspect "*place*" should be identified with the inter-aspect relations.



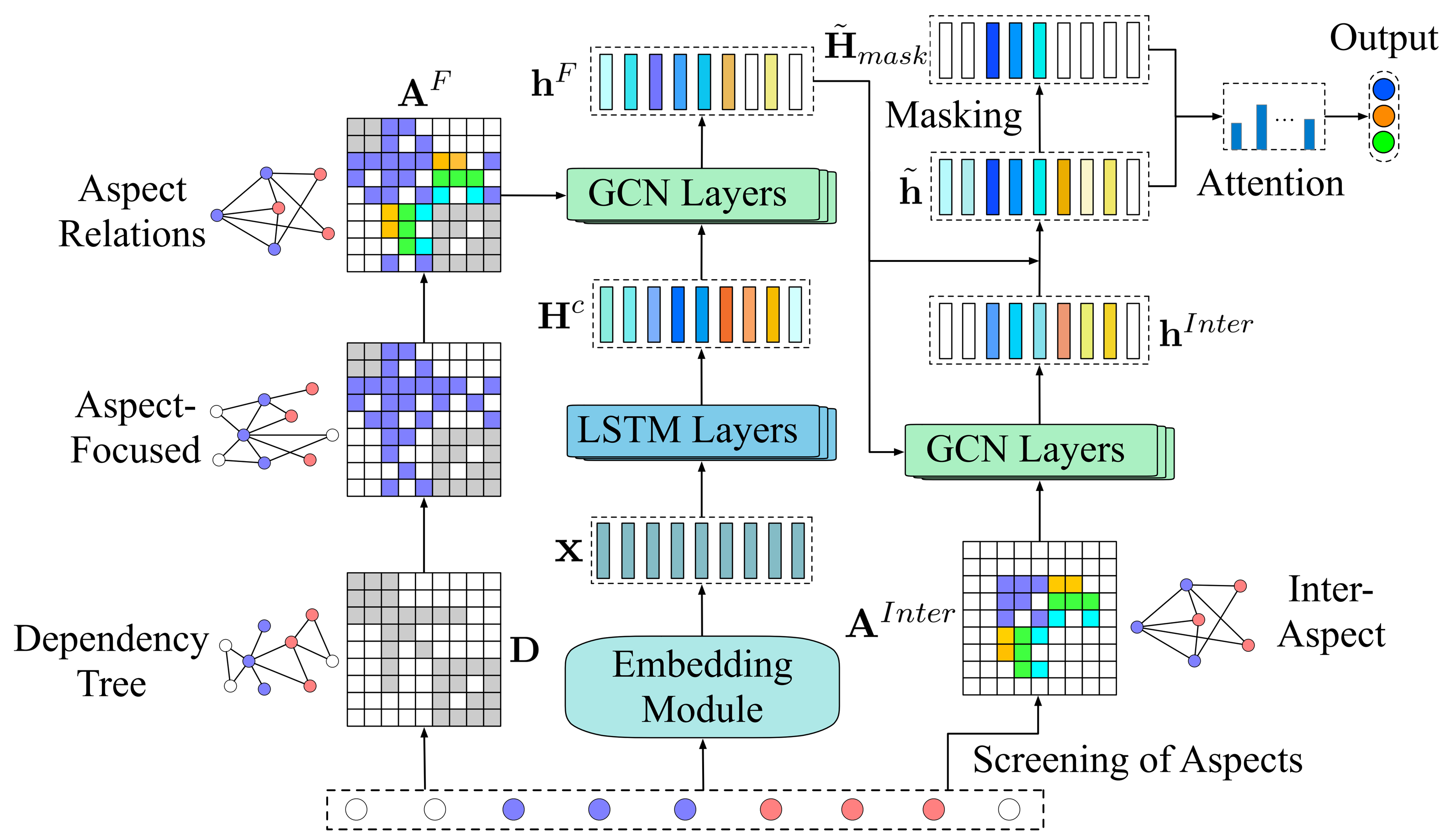
(a) Example of aspect-focused relations (b) Example of inter-aspect relations

Key Contributions:

- We explore a novel solution to construct the graph for each instance, in which both aspect-focused and inter-aspect syntactical dependencies are introduced.
- An Interactive GCN model is proposed to derive aspect-specific sentiment features by interactively extracting the sentiment relations within aspect words and across different aspects in the context.
- Experimental results on four benchmark datasets show that the proposed model achieves the state-of-the-art performance in aspect sentiment analysis.

Method

Framework Architecture of the Proposed Interactive GCN:



Producing Ordinary Graphs over Dependency Tree:

We first produce an ordinary dependency graph $\mathbf{D} \in \mathbb{R}^{n \times n}$ for each input sentence over the dependency tree with spaCy toolkit:

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

Interactive Graph Convolutional Network:

Each node in the l-th GCN layer is updated as follow:

$$\mathbf{h}_i^l = \mathrm{ReLU}(\tilde{\mathbf{A}}_i^F \mathbf{g}_i^{l-1} \mathbf{W}^l + \mathbf{b}^l) \tag{2}$$

where $\mathbf{g}_i^{l-1}$ is the hidden representation of the preceding GCN layer. Afterward, we combine the final representations of aspect-focused and inter-aspect features:

$$\tilde{\mathbf{h}} = \mathbf{h}^F + \gamma \mathbf{h}^{Inter} \tag{3}$$

where γ is the coefficient of inter-aspect features.

Constructing Graphs for Aspect-focused and Inter-aspect:

We refine the graph via incorporating a relative position weight according to the specific aspect:

$$W_{i,j}^F = \begin{cases} 1 & \text{if } w_i \in \{a_i^s\} \text{ and } w_j \in \{a_i^s\} \\ 1/(|j - p^s| + 1) & \text{if } w_i \in \{a_i^s\} \\ 1/(|i - p^s| + 1) & \text{if } w_j \in \{a_i^s\} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $|\cdot|$ is an absolute value function, p^s is the beginning position of the specific aspect, $\{a_i^s\}$ is the word set of the specific aspect. The aspect-focused syntactical dependency adjacency matrix can be derived as:

$$G_{i,j} = \begin{cases} D_{i,j} + D_{i,j} * W_{i,j}^F & \text{if } D_{i,j} = 1 \\ W_{i,j}^F & \text{otherwise} \end{cases} \tag{5}$$

We further refine the aspect-focused graph via incorporating relative graphs from other aspects into the aspect-focused adjacency matrix:

$$\mathbf{A}^F = \mathbf{G} + \frac{1}{l} \sum_{a \in \{a_i^o\}} (\alpha * \mathbf{G}_a), \quad \alpha = 1/(|p^o - p^s| + 1) \tag{6}$$

where $\{a_i^o\}$ is the word set of length l of other aspects, and the p^o for each $a \in \{a_i^o\}$ denotes the beginning position of the other aspect. We screen the aspects from the sentence and construct an inter-aspect adjacency matrix for these aspects to derive the contextual sentiment dependencies of these aspects:

$$A_{i,j}^{Inter} = \begin{cases} 1 + 1/(|j - p^s| + 1) & \text{if } w_i \in \{a_i^s\} \text{ and } w_j \in \{a_i^o\} \\ 1 + 1/(|i - p^s| + 1) & \text{if } w_j \in \{a_i^s\} \text{ and } w_i \in \{a_i^o\} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

To enrich the information of dependencies for the input sentence, we construct the adjacency matrices in un-directional, i.e. $A_{i,j}^F = A_{j,i}^F$ and $A_{i,j}^{Inter} = A_{j,i}^{Inter}$.

Experimental Results & Conclusion

Comparison Methods:

IACapsNet (Du et al., 2019), **TD-GAT** (Huang and Carley, 2019), **ASGCN** (Zhang et al., 2019), **BERT** (Devlin et al., 2019), **SA-GCN+BERT** (Hou et al., 2019).

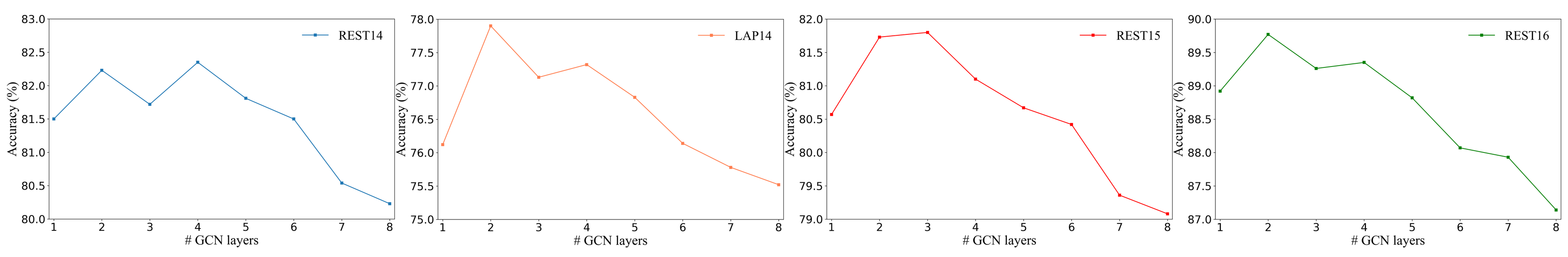
Main Experimental Results:

Model	REST14 (%)		LAP14 (%)		REST15 (%)		REST16 (%)	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
IACapsNet	81.79	73.40	76.80	73.29	-	-	-	-
TD-GAT	81.10	-	73.70	-	-	-	-	-
ASGCN	80.77	72.02	75.55	71.05	79.89	61.89	88.99	67.48
BERT	84.11	76.68	77.59	73.28	83.48	66.18	90.10	74.16
SA-GCN+BERT	85.80	79.70	81.70	78.80	-	-	-	-
InterGCN (ours)	**82.23**	**74.01**	**77.86**	**74.32**	**81.76**	**65.67**	**89.77**	**73.05**
InterGCN+BERT (ours)	**87.12**	**81.02**	**82.87**	**79.32**	**85.42**	**71.05**	**91.27**	**78.32**

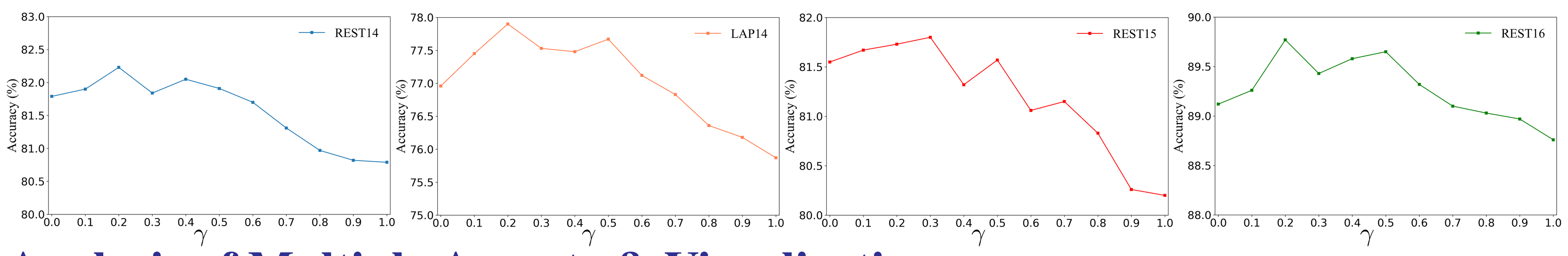
Experimental Results of Ablation Study:

Model	REST14 (%)		LAP14 (%)		REST15 (%)		REST16 (%)	
	Acc.	F1	Acc.	F1	Acc.	F1	Acc.	F1
InterGCN w/o dependency	82.03	73.82	76.70	72.63	81.56	65.12	89.43	72.67
InterGCN w/o aspect-focused	81.32	72.57	75.89	71.45	80.06	62.73	89.05	68.36
InterGCN w/o aspect relations	81.80	73.61	76.38	72.21	81.34	64.57	89.21	71.89
InterGCN w/o inter-aspect	81.79	73.42	76.96	73.29	81.55	65.08	89.12	70.60
InterGCN	**82.23**	**74.01**	**77.86**	**74.32**	**81.76**	**65.67**	**89.77**	**73.05**

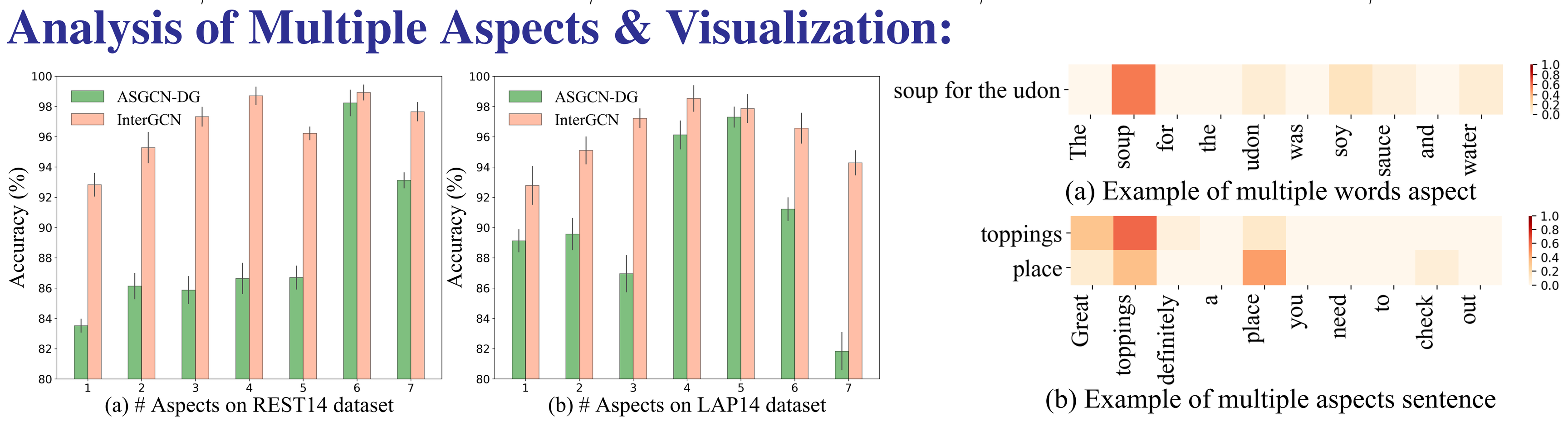
The Impact of GCN Layers:



The Impact of the Coefficient of Inter-Aspect:



Analysis of Multiple Aspects & Visualization:



Conclusion:

We explore a novel solution of constructing aspect-focused and inter-aspect dependency graphs for aspect sentiment analysis. Based on it, an Interactive Graph Convolutional Networks (InterGCN) model is proposed to extract the aspect-specific sentiment features from the aspect-focused and inter-aspect perspective.