



Research Article

Interactive Attention-Based Convolutional GRU for Aspect Level Sentiment Analysis

Lisha Chen*, Tianrui Li, Huaishao Luo, Chengfeng Yin

School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu, China

ARTICLE INFO

Article History

Received 20 March 2021

Accepted 04 July 2021

Keywords

Sentiment classification
convolutional neural network
gated recurrent units
attention mechanism

ABSTRACT

Aspect level sentiment analysis aims at identifying sentiment polarity towards specific aspect terms in a given sentence. Most methods based on deep learning integrate Recurrent Neural Network (RNN) and its variants with the attention mechanism to model the influence of different context words on sentiment polarity. In recent research, Convolutional Neural Network (CNN) and gating mechanism are introduced to obtain complex semantic representation. However, existing methods have not realized the importance of sufficiently combining the sequence modeling ability of RNN with the high-dimensional feature extraction ability of CNN. Targeting this problem, we propose a novel solution named Interactive Attention-based Convolutional Bidirectional Gated Recurrent Unit (IAC-GRU). IAC-GRU not only incorporates the sequence feature extracted by Bi-GRU into CNN to accurately predict the sentiment polarity, but also models the target and the context words separately and learns mutual influence between them. Additionally, we also incorporate the position information and Part-of-Speech (POS) information as prior knowledge into the embedding layer. The experimental results on SemEval2014 datasets show the effectiveness of our proposed model.

© 2021 The Authors. Publishing services by Atlantis Press International B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

1. INTRODUCTION

Sentiment analysis is a vital task in Natural Language Processing (NLP), which has received extensive attention from academic and industry. Traditional sentiment analysis can be divided into two categories: the document level and the sentence level, assuming that a document or a sentence contains only one sentiment polarity, respectively. The above assumption is unreasonable in many cases, which is compensated by fine-grained aspect-level sentiment analysis. Aspect-level sentiment analysis aims at identifying sentiment polarity towards different aspect terms in a given sentence. For instance, in the sentence “the bathroom is tidy but the service is terrible.”, the polarity for the aspect “bathroom” and “service” are positive and negative separately.

Traditional approaches generally design rich features for sentiment classification [1,2], but the method is time-consuming and labor-intensive. More importantly, its performance greatly depends on the quality of feature engineering and has encountered bottlenecks. With the great success of deep learning in machine translation [3], question answering [4], and other fields, researchers have introduced Recurrent Neural Network (RNN) and its variants into aspect level sentiment analysis, such as target-dependent Long-Short Term Memory (TD-LSTM) and target-connection LSTM (TC-LSTM) proposed by Tang et al. [5] which models the interaction between target and context words, and Hierarchical Bidirectional LSTM proposed by Ruder et al. [6]. To enforce the model to pay much attention to the aspect-specific words, attention mechanism

have been introduced into deep neural network and achieved outstanding performance, e.g. ATAE-LSTM [7], IAN [8]. After incorporating various additional auxiliary information like location information, commonsense information, etc. into the model, RNN-based model has already achieved advanced performance while it's complex and time-consuming during training. Gated mechanism-based Convolutional Neural Network (CNN) provides researchers alternative approach with simpler structure and superior performance [9-11].

In previous work, either a complex RNN-based attention mechanism or gated mechanism-based CNN was used to improve performance. But both approaches sacrifice the advantages of the other. Therefore, this paper proposes a neural network model, named Interactive Attention-based Convolutional Bidirectional Gated Recurrent Unit (IAC-GRU), which integrates CNN and Bi-GRU to effectively extract complex semantics and sequence features as well as aspect-specific correlation. In addition, position and Part-of-Speech (POS) information are incorporated into word embedding to distinguish the crucial information of the context words for polarity detection. Inspired by previous studies, an interactive attention network is used to model targets and context words respectively and calculate their mutual effects.

The contributions of this paper is as follows: (1) A novel IAC-GRU based on the fusion of CNN and Bi-GRU is presented. (2) The effectiveness of the position information of context words relative to aspect terms is investigated. (3) The validity of the filtered POS information of context words and target words is studied. (4) Experimental results indicate that our IAC-GRU surpasses the state-of-art baselines on the widely used SemEval2014 datasets.

* Corresponding author. Email: lschen@my.swjtu.edu.cn

Peer review under responsibility of KEO (Henan) Education Technology Co. Ltd

The remaining of this paper is organized as follows. Section 2 discusses the existing works in aspect level sentiment analysis. Section 3 details the proposed model including word embedding, neural network and attention mechanism. Section 4 shows the competitive results on well-known datasets and verifies the effectiveness of our model. Finally, we summarize our work in Section 5.

2. RELATED WORK

Different from traditional sentiment analysis, aspect level sentiment analysis aims at identifying sentiment polarity towards different aspect terms in a given sentence. Earlier methods are generally based on dictionary or machine learning depended on various kinds of artificial features. Dictionary-based method generally aggregates the sentiment features of surrounding words to aspect term itself. In [12], Hu and Liu assigned a sentiment label to each adjective in the sentence by looking up the sentiment dictionary. If there are negators in the surrounding five words, the sentiment polarity will be reversed. Finally, the majority voting is used to assign the same sentiment labels to all targets in the sentence. Traditional methods based on machine learning are commonly achieved by constructing semantic features to train classifiers. Yu and Zha [13] employed a parse tree to obtain sentiment words for each aspect and used SVM to obtain the sentiment label. Jiang et al. [14] explored the grammar structures to build SVM classifier. Rao et al. [15] applied a graph-based semi-supervised learning framework to build sentiment lexicons, which has achieved competitive performance. However, aforementioned methods are time-consuming and labor-intensive.

Recently, a variety of deep neural networks have been proposed and achieved remarkable results in aspect level sentiment analysis, e.g. Recursive Neural Network [16–18], Recurrent Neural Network [5–7], Convolutional Neural Network [19], Memory Network [20–23].

Tang et al. [5] first introduced LSTM into aspect level sentiment analysis to capture the relationship between target and context words. Jiang et al. [1] revealed that 40% of the errors were due to neglected target information. Hence, Tang et al. [5] divided a sentence into left and right parts fused with target information and used two separate LSTMs to model the relation. LSTM is limited to capturing the information transmitted from the proceeding text to the following text. Ruder et al. [6] applied a hierarchical bidirectional LSTM to leverage the intra- and inter-sentence relationships. To enforce the model to attend important information related to targets, Wang et al. [7] took the concatenations of the word embedding and aspect embedding as the input of LSTM and joined the attention mechanism behind the hidden layer to classify the sentiment label which validate the effectiveness of attention mechanism. The biggest performance limitation of fine-grained sentiment analysis lies in the serious lack of training samples. He et al. [24] adopted pre-training and multi-task learning based on attentive LSTM model to transfer the knowledge learned from extensive document level corpus to aspect level sentiment analysis. Ma et al. [8] made use of LSTM to extract sequence features of target and context words respectively to monitor the generation of attention weight vector. Compared with LSTM, CNN can capture higher level semantic features including local and global features.

Huang et al. [9] introduced a parameterized CNN. Xue et al. [19] incorporated the gating mechanism into CNN with multiple filters to extract n -gram features. Their performance proved that CNN-based method is a potential direction worth exploring. Memory network has achieved great success in the field of natural language processing and provides a new solution for aspect level sentiment analysis. Tang et al. [22] first proposed a multi-hop attention model based on external storage and used multiple computing layers to calculate the importance of text features. Li et al. [23] also presented an end-to-end method based on deep memory network to jointly perform aspect extraction and aspect-based sentiment analysis.

3. PROPOSED MODEL

In this section, we start with an overview of the proposed IAC-GRU model. Then, each layer of the IAC-GRU model is introduced in detail. Finally, we introduce the loss function. For aspect level sentiment analysis, we assume that the sentence, $S = \{w_1, w_2, \dots, w_N\}$, the aspect term, $T = \{t_1, t_2, \dots, t_M\}$ in which N represents the length of sentence, M represents the length of target and T is a subsequence of S . The goal of IAC-GRU model is to infer the sentiment polarity of different targets in a given sentence. Firstly, the target and context words are encoded respectively and position information and POS information are fused in word embedding. Next, word embeddings are input into separate two Bi-GRUs to collect the sequence features in a sentence. Then, a hidden layer state from Bi-GRU is fed into CNN and the maximum pooling is employed to calculate the interaction between target and context words. Finally, the output of two attention mechanisms is concatenated for a classifier. The overall architecture of our model is shown in Figure 1.

3.1. Position Embedding

Earlier studies have realized the significance of position information, and generally used pre-defined strategy and training as model parameters to model it. Tang et al. [22] combined two ways of taking position information into consideration and introduced them into the field of aspect level sentiment analysis. Inspired by previous approaches, we believe that context word closer to the target contributes more to the sentiment polarity given target. In order to quantify the correlation between target and context words, we use a relative distance to obtain a one-dimensional word position index or a word position matrix according to one or more aspect terms appearing in the sentence. We assume that the target itself in word position index is marked as 0. The relative distance of words r_i is calculated as in (1):

$$r_i = \begin{cases} i - t_e, & i > t_e \\ 0, & t_e \leq i \leq t_s \\ t_s - i, & i < t_s \end{cases} \quad (1)$$

where r_i indicates the relative distance of i -th words in a sentence; t_s and t_e indicate the beginning and end of the target, respectively. For giving a more vivid explanation, we use an example below to illustrate it. Given a sentence “All the pastas are fantastic and the homemade lasagna is some of the best that I have had in the City.”, a position index [2, 1, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19] will be gained for a specific target “pastas”. Obviously, the position index of aspect term is assigned to 0, and other context words’ index increases with the growing distance. When we set

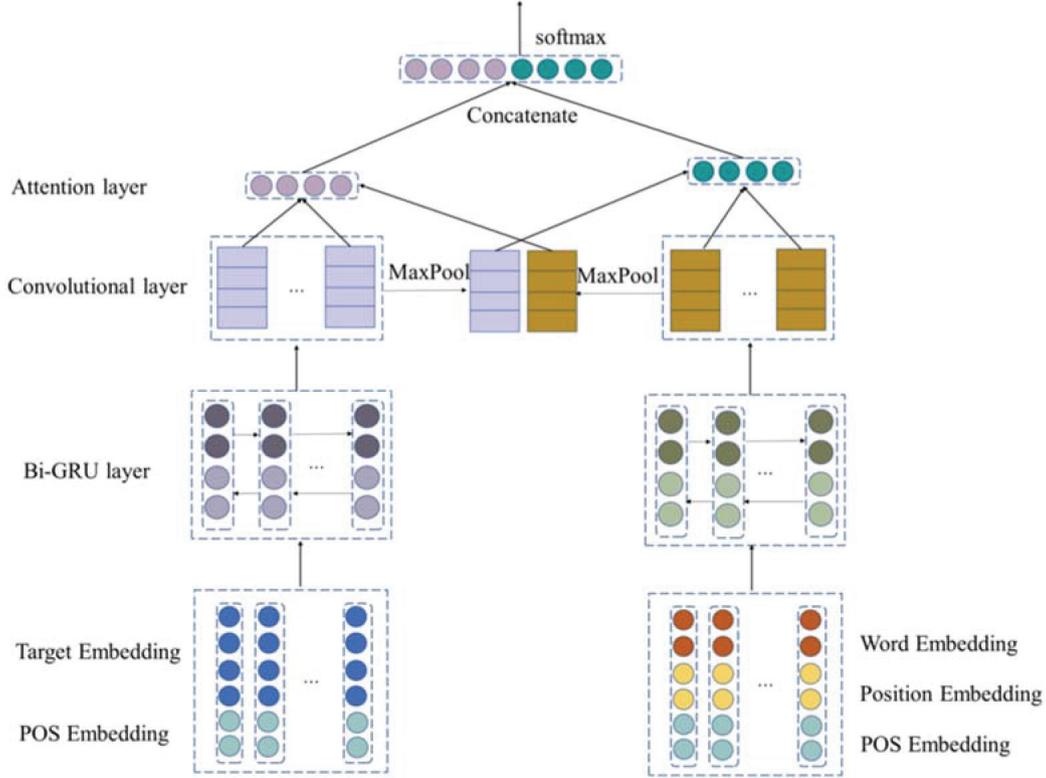


Figure 1 | Architecture of the proposed model for aspect level sentiment analysis.

“homemade lasagna” as center aspect term, the position index list will be [7, 6, 5, 4, 3, 2, 1, 0, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. After obtaining the relative distance r_i , we calculate the position embedding $p = [p_1, p_2, \dots, p_N]$ by looking up the position matrix $P \in R^{d^p \times N}$, which is randomly initialized and updated during the training process, where d^p represents the dimension of position embedding, and N represents the length of the sentence.

3.2. POS Embedding

POS information is drawn as a prior knowledge into sentiment analysis to infer the sentiment polarity. However, not all POS categories have contribution to sentiment polarity, and blindly using POS categories introduces additional noise into the proposed model. Inspired by Shuang et al. [25], we assign POS tags to each word in the sentence from the candidate set $\{P_{adj}, P_{adv}, P_{verb}, P_{other}\}$. Taking the sentence “We recently decided to try this location, and to our delight, they have outdoor seating, perfect since I had my yorkie with me.” as an example, the natural language toolkit (NLTK) is employed to obtain the POS index in our experiment. Thus the POS index of sentence is represented as $pos^c = \{P_{other}, P_{adv}, P_{other}, P_{other}, P_{verb}, P_{other}, P_{other}\}$. The POS index of aspect is represented as $pos^t = \{P_{other}\}$. We acquire the final POS embedding of target $pos^t = [pos_1^t, pos_2^t, \dots, pos_M^t]$ and context words $pos^c = [pos_1^c, pos_2^c, \dots, pos_N^c]$ respectively by looking up the POS matrix $P_{pos} \in R^{d^{pos} \times Num}$, which is randomly initialized and updated during the training process, where d^{pos} indicates the dimension of POS embedding, and Num indicates the number of POS tags.

3.3. Word Embedding

We map the context of N words $[w_1, w_2, \dots, w_N]$ and the target of M words $[t_1, t_2, \dots, t_M]$ into a low-dimensional dense continuous embedding by looking up the embedding matrix $E \in R^{v \times d}$, where v indicates the vocabulary size and d indicates the dimension of word embedding. In our experiment, E is initialized by pre-trained word embedding vector obtained by Glove [29]. We label the context embedding as $[e_1^c, e_2^c, \dots, e_N^c]$ and label the aspect embedding as $[e_1^t, e_2^t, \dots, e_M^t]$, separately. For the i -th word, we concatenate word embedding and position embedding and POS embedding as $x_i^c = e_i^c \oplus p_i \oplus pos_i^c$. Likewise, on account of the position index of target itself is marked as 0, we simply connect word embedding and POS embedding of target labeled as $x_i^t = e_i^t \oplus pos_i^t$.

3.4. Bidirectional Gated Recurrent Units

Standard RNN [26] has achieved excellent performance in numerous NLP tasks while it suffers from the problems of gradient explosion and vanishing gradient. LSTM [27], as a variant used to overcome the defects of RNN, does well in learning long-term dependencies. GRU [28] is similar to LSTM along with a simpler structure. Instead of introducing additional memory units, GRU adds an update gate based on the original RNN to control how much information the current state needs to retain from the historical state and the degree of receiving from the candidate state. GRU also uses a reset gate to control the influence of the previous state on the current state. The GRU state is calculated as follows:

$$z_t = \delta(W_z x_t + U_z h_{t-1} + b_z) \quad (2)$$

$$r_t = \delta(W_r x_t + U_r h_{t-1} + b_r) \quad (3)$$

$$\tilde{h}_t = \tanh(W_h x_t + r_{t-1} \odot (U_h h_{t-2}) + b_h) \quad (4)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (5)$$

where z_t represents the update gate, \odot represents element-wise product, r_t represents the reset gate and \tilde{h}_t represents the candidate state. In order to capture the information from backward text to forward text based on GRU, we employ Bi-GRU to acquire the context information of complete time series in sentences. We feed the aspect embedding into left Bi-GRU and connect the hidden state of forward GRU $\vec{h}_i^t \in R^{d_h}$ and backward GRU $\overleftarrow{h}_i^t \in R^{d_h}$ as the final hidden state $h_i^t = [\vec{h}_i^t; \overleftarrow{h}_i^t] \in R^{2d_h}$, where d_h indicates the number of hidden units. Similarly, we take the context word embedding as the input of right Bi-GRU to obtain the final hidden state $h_i^c = [\vec{h}_i^c; \overleftarrow{h}_i^c] \in R^{2d_h}$.

3.5. Convolutional Neural Network

We stitch CNN behind the hidden layer state of Bi-GRU to extract the local and global semantic features inside the targets and inside the context. For context words, we take the right hidden state as input and use the convolutional layer with multiple convolutional kernels along with different widths. We slide a convolution filter $W_i^c \in R^{K_i^c \times 2d_h}$ on the input window to obtain new features. Then, we use a relu function and a max-pooling layer to obtain the most significant features of context words g_i^c as in (6). For the purpose of computing the interaction between targets and context words, we connect the output of q convolution filters as final influence of context to aspect labeled as $I^c = [g_1^c \oplus g_2^c \oplus \dots \oplus g_q^c]$. Likewise, we input the hidden layer state of the target to left CNN to extract the most important features $I^t = [u_1^t \oplus u_2^t \oplus \dots \oplus u_p^t]$ of p convolution filters, in which u_i^t is calculated as in (7).

$$g_i^c = \text{Max} \left(\text{relu} \left(h_{i:i+K_i^c}^c * W_i^c + b_c \right) \right) \quad (6)$$

$$u_i^t = \text{Max} \left(\text{relu} \left(h_{i:i+K_i^t}^t * W_i^t + b_t \right) \right) \quad (7)$$

where $*$ represents the convolutional operation, K_i^c and K_i^t are the width of i -th context convolution filter and i -th target convolution filter respectively; $W_i^t \in R^{K_i^t \times 2d_h}$ is the target convolution filter; $b_c \in R$ and $b_t \in R$ are bias, respectively.

3.6. Interactive Attention Network

Each word in a sentence has a different contribution to sentiment polarity. In order to enforce the proposed model pay attention to words more relevant to a given aspect, we introduce an interactive attention mechanism to treat the target and context words equally and calculate their interaction to supervise the generation of attention weight vector. We obtain the attention weights vector of aspect to context words α_i and context to aspect words β_i as follows:

$$\alpha_i = \frac{\exp(\mu(g_i^c, I^t))}{\sum_{j=1}^q \exp(\mu(g_j^c, I^t))} \quad (8)$$

$$\mu(g_i^c, I^t) = \tanh(g_i^c W_a (I^t)^T + b_a) \quad (9)$$

$$\beta_i = \frac{\exp(\mu(u_i^t, I^c))}{\sum_{j=1}^p \exp(\mu(u_j^t, I^c))} \quad (10)$$

where W_a and b_a are weight matrix and bias, respectively. On the basis of the calculated attention weights, the final aspect representation v and context representation a are as follows respectively:

$$v = \sum_{i=1}^q \alpha_i g_i^c \quad (11)$$

$$a = \sum_{i=1}^p \beta_i u_i^t \quad (12)$$

We concatenate the aspect representation v and context representation a as final sentence representation for a classifier labeled as S and append a non-linear layer as in (13).

$$Z = \tanh(W_s S + b_s) \quad (13)$$

Finally, we use a softmax layer to obtain the probability distribution of sentiment polarity as positive, negative and neutral as in (14).

$$p = \text{soft max}(W_z Z + b_z) \quad (14)$$

3.7. Model Training

The objective of model optimization in our paper is to minimize cross-entropy loss between y and \tilde{y} for the training dataset. We assume that y is the true value and the \tilde{y} is the predict value. The $L2$ regularization is employed for cross-entropy. Let λ be the coefficient for $L2$ regularization and θ contain all the parameters.

$$l = - \sum_a \sum_b \gamma_a^b \log \tilde{y}_a^b + \lambda \|\theta\|^2 \quad (15)$$

4. EXPERIMENTS

4.1. Datasets

We conduct our experiments on SemEval 2014 Task 4, which contains abundant customer reviews of Restaurants and Laptop domains. The source datasets are signed with four sentiment polarity: positive, negative, conflict and neutral. Drawing on the experience of existing approaches, we remove conflict polarity. The details of data set are shown in Table 1.

4.2. Experimental Setting

In our experiments, we employ 300-dimensional word embedding vectors pre-trained by Glove [29]. Both position embedding dimension and POS dimension as auxiliary information are set to 100 and 36 as in [31], which is initialized randomly and fine-tuned during training process. Hidden units are set to 200. The coefficient for $L2$ regularization is set to $1e-5$ and dropout is set to 0.5 as in [8]. In the relevant parameter settings of CNN, the kernel num is set to 100. The kernel sizes for context words and target terms are set to 1, 3,

Table 1 | Details of the SemEval2014

DataSet	Laptop			Restaurants		
	Positive	Neural	Negative	Positive	Neural	Negative
Train	987	460	866	2164	633	805
Test	341	169	128	728	196	196

5 and 1 as in [31], respectively. The batch size is set to 32 and the num of epoch is set to 100. We select Adam as an optimizer with a learning rate of 0.001. We choose Accuracy as the evaluation metric.

4.3. Baselines

In order to validate the effectiveness of model, we compare our proposed model with the following several classic and competitive models:

TD-LSTM: It uses a left one LSTM_L and a right one LSTM_R to model the preceding and following contexts separately to take the target information into consideration and concatenates the hidden states from two LSTMs for final prediction [5].

ATAE-LSTM: It feeds aspect embedding together with word embedding into LSTM and uses the hidden states appended with aspect embedding to generate the attention weight vector [7].

IAN: Interactive attention network (IAN) [8] employs two separate LSTMs to treat target words and context words equally, and uses the average-pooling of the hidden states to supervise the interactive attention weight vector.

GCAE: Gated convolutional network with aspect embedding (GCAE) [19] uses a convolutional filter to capture n -gram features and utilizes the gating units and max-pooling layer to extract sentiment feature.

CMA-MemNet: Multi-head self-attention memory network (CMA-MemNet) [30] uses a memory network to extract rich semantic information in sequences and combines a convolutional network with self-attention to make up for the grammatical information and the semantic information of the sequence itself ignored by memory network.

IGCN: Compared with other CNN-based network, IGCN [31] takes the separate modeling of targets into account and uses gate mechanism to learn their interaction. It also incorporates position information and POS information.

4.4. Results

As shown in Table 2, the proposed IAC-GRU model achieves the best performance compared with baseline models on both Restaurants and Laptop datasets. TD-LSTM performs the worst, although it utilizes the standard LSTM to capture target-dependent feature information in preceding and following context. It still doesn't pay enough attention to the target information. ATAE-LSTM obtains

better performance than TD-LSTM, as it employs attention mechanism to enforce the model to show solicitude for the most target-relevant information in a given sentence. IAN uses an interactive attention network to capture the interaction between aspect and context words, thus it performs better than TD-LSTM and ATAE-LSTM. CGAE introduces a convolutional network and gate mechanism to gain n -gram aspect-relevant features at multi-granularity. Therefore, its performance is reasonably superior than previous LSTM-based models. CMA-MemNet simultaneously uses CNN and memory network to learn the rich semantic features as well as the syntactic features related to targets. CMA-MemNet has made competitive improvement than CGAE. IGCN has the similar performance as CMA-MemNet on Restaurants datasets due to its independent modeling of targets and interactive learning, but it obtains higher accuracy than CMA-MemNet on Laptop datasets. IAC-GRU feeds position and POS information as auxiliary information together with word embedding into Bi-GRU which has similar performance and simpler structure compared with LSTM to obtain significant sequence feature. CNN is also applied to extract high-level semantic features and integrate interactive attention networks. Hence IAC-GRU outperforms all baseline methods on Restaurants and Laptop datasets.

4.5. Case Study

We visualize the attention weight in several samples as Figure 2. The details of samples are indicated in Table 3. We randomly select two sentences from the Restaurants test dataset which contains a single and multiple aspect terms, respectively. In the first example, the neighboring words of the target "people" do not contain corresponding sentiment words. Even so, the model still pays attention to the negative words "never" and the verb "disappoints" that describes target "people". This approves that our IAC-GRU model can capture the context words which play a vital role in inferring the sentiment polarity. In Case 2, when aspect term is "waiter", the affective word "friendly" which is close to each other gets higher attention score, while far away from each other such as "shame", "supportive" gets less attention. This fully gives sufficient expression to the significance of position information. In view of the sentiment word "supportive", the aspect term "staff" is easy to be judged as positive. Nevertheless, its real sentiment polarity needs to be

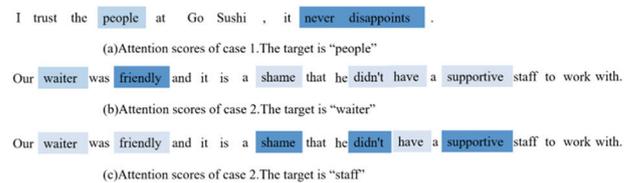


Figure 2 | The attention scores proposed by our model. Deeper color marks higher attention score.

Table 2 | Experimental results for aspect-level sentiment analysis

	Restaurant	Laptop
TD-LSTM [22]	75.63	68.13
ATAE-LSTM [7]	77.2	68.7
IAN [8]	78.60	72.10
CGAE [19]	77.28	69.14
CMA-MemNet [30]	81.26	73.24
IGCN [31]	81.34	75.24
IAC-GRU _{w/o} position	80.63	73.20
IAC-GRU _{w/o} POS	81.24	73.69
IA-GRU	81.07	75.01
IAC-GRU	82.32	76.02

Table 3 | Details of the samples

Description	Text
Single target (case 1)	I trust the [people] _{positive} at Go Sushi, it never disappoints.
Multi-targets (case 2)	Our [waiter] _{positive} was friendly and it is a shame that he didn't have a supportive [staff] _{negative} to work with.

obtained by combining the negator “didn’t” or the distant adjective “shame”. When the sentence contains multiple targets or negative words, our IAC-GRU model still obtains correct result. This means that IAC-GRU fully learns the internal representation of the aspect term and the interaction of word embedding and position and POS information between target and context words.

4.6. Ablation Study

To investigate the effectiveness of CNN, position information, and POS information, we compare IAC-GRU with three ablated models, namely, IA-GRU which only removes the CNN component, IAC-GRU_{w/o position} which only removes the position embedding, and IAC-GRU_{w/o POS} which only removes the POS embedding. The results are shown in Table 2. Obviously, IAC-GRU_{w/o position} decreases by 1.69% and 2.82% on Restaurants and Laptop datasets, which indicates that the position information incorporated in the model will make the words close to aspect get higher weight as the auxiliary information of filtering classification. The performance of IAC-GRU_{w/o POS} drops 1.08% and 2.33% compared with the complete model, which indicates that words marked as different POS categories will be treated differently during model training. IA-GRU only uses Bi-GRU to extract sequence information and ignores context information besides word level. Hence, its performance is inferior to the complete model. In summary, IAC-GRU outperforms three ablated models, which reconfirms that the three critical components have significant impact on performance.

5. CONCLUSIONS

In this paper, we proposed an interactive attention-based convolutional GRU for aspect-level sentiment analysis (IAC-GRU). IAC-GRU incorporates the position and POS information into word embedding and uses Bi-GRU and CNN to capture the sequence feature and the high-level local and global features simultaneously. Additionally, we constructed an interactive attention network to treat the target and context words equally and model the mutual interaction. The experimental results on SemEval2014 datasets (i.e., Restaurants and Laptop datasets) demonstrated the effectiveness of the proposed IAC-GRU model. Finally, we designed three ablation study with CNN, position embedding, and POS embedding as the unique variable. The descending performance demonstrated the effectiveness of the three critical components.

Finally, in this paper, we mainly focus on the performance of position information and POS information in this task, while the interaction between aspects is not explicitly considered. In the future, we will explore the interaction between multiple aspects and introduce more efficient and concise pre-trained models such as BERT to improve our task performance.

CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest.

AUTHORS' CONTRIBUTION

Lisha Chen contributed in methodology, software, validation and writing. Tianrui Li, Huaishao Luo and Chengfeng Yin contributed in review & editing.

ACKNOWLEDGMENT

This work was supported by the Key R & D Project of Sichuan Province, China (No. 2020YFG0035).

REFERENCES

- [1] L. Jiang, M. Yu, M. Zhou, X. Liu, T. Zhao, Target-dependent twitter sentiment classification, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, ACM, Portland, USA, 2011, pp. 151–160.
- [2] S. Kiritchenko, X. Zhu, C. Cherry, S. Mohammad, NRC-Canada-2014: detecting aspects and sentiment in customer reviews, Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval), Association for Computational Linguistics, Dublin, Ireland, 2014, pp. 437–442.
- [3] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, Learning phrase representation using RNN encoder-decoder for statistical machine translation, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Doha, Qatar, 2014, pp. 1724–1734.
- [4] J. Andreas, M. Rohrbach, T. Darrell, D. Klein, Learning to compose neural networks for question answering, Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, San Diego, California, 2016, pp. 1545–1554.
- [5] D. Tang, B. Qin, X. Feng, T. Liu, Effective LSTMs for target-dependent sentiment classification, Proceedings of COLING 26th International Conference on Computational Linguistics: Technical Papers, The COLING 2016 Organizing Committee, Osaka, Japan, 2016, pp. 3298–3307.
- [6] S. Ruder, P. Ghaffari, J.G. Breslin, A hierarchical model of reviews for aspect-based sentiment analysis, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Austin, Texas, 2016, pp. 999–1005.
- [7] Y. Wang, M. Huang, X. Zhu, L. Zhao, Attention-based LSTM for aspect-level sentiment classification, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Austin, Texas, 2016, pp. 606–615.
- [8] D. Ma, S. Li, X. Zhang, H. Wang, Interactive attention networks for aspect-Level sentiment classification, Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI, Melbourne, Australia, 2017, pp. 4068–4074.
- [9] B. Huang, K. Carley, Parameterized convolutional neural networks for aspect level sentiment classification, Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 1091–1096.
- [10] C. Fan, Q. Gao, J. Du, G. Lin, R. Xu, K.F. Wong, Convolution-based memory network for aspect-based sentiment analysis, Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 2018, pp. 1161–1164.
- [11] X. Li, L. Bing, W. Lam, B. Shi, Transformation networks for target-oriented sentiment classification, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 946–956.
- [12] M. Hu, B. Liu, Mining and summarizing customer reviews, Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Seattle, USA, 2004, pp. 168–177.
- [13] J. Yu, Z.J. Zha, M. Wang, T.S. Chua, Aspect ranking: identifying important product aspects from online consumer reviews, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Portland, OR, USA, 2011, pp. 1496–1505.

- [14] L. Jiang, M. Yu, M. Zhou, X. Liu, T. Zhao, Target-dependent twitter sentiment classification, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, Portland, OR, USA, 2011, pp. 151–160.
- [15] D. Rao, D. Ravichandran, Semi-supervised polarity lexicon introduction, Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics, ACM, Athens, Greece, 2009, pp. 675–682.
- [16] R. Socher, J. Pennington, E.H. Huang, A.Y. Ng, C.D. Manning, Semi-supervised recursive autoencoders for predicting sentiment distributions, Proceedings of the Conference on Empirical Methods in Natural Language Processing, ACL, Stroudsburg, PA, USA, 2011, pp. 151–161.
- [17] L. Dong, F. Wei, C. Tan, D. Tang, M. Zhou, K. Xu, Adaptive recursive neural network for target-dependent twitter sentiment classification, Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Baltimore, Maryland, 2014, pp. 49–54.
- [18] T.H. Nguyen, K. Shirai, PhraseRNN: phrase recursive neural network for aspect-based sentiment analysis, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Lisbon, Portugal, 2015, pp. 2509–2514.
- [19] W. Xue, T. Li, Aspect based sentiment analysis with gated convolutional networks, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 2514–2523.
- [20] Y. Tay, L.A. Tuan, S.C. Hui, Dyadic memory networks for aspect-based sentiment analysis, Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, ACM, 2017, pp. 107–116.
- [21] P. Zhu, T. Qian, Enhanced aspect level sentiment classification with auxiliary memory, Proceedings of the 27th International Conference on Computational Linguistics, Association for Computational Linguistics, Santa Fe, New Mexico, USA, 2018, pp. 1077–1087.
- [22] D. Tang, B. Qin, T. Liu, Aspect level sentiment classification with deep memory network, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Austin, Texas, 2016, pp. 214–224.
- [23] X. Li, W. Lam, Deep multi-task learning for aspect term extraction with memory interaction, Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 2886–2892.
- [24] R. He, W.S. Lee, H.T. Ng, D. Dahlmeier, Exploiting document knowledge for aspect-level sentiment classification, Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Melbourne, Australia, 2018, pp. 579–585.
- [25] K. Shuang, M. Gu, R. Li, J. Loo, S. Su, Interactive POS-aware network for aspect-level sentiment classification, Neurocomputing 420 (2021), 181–196.
- [26] J.L. Elman, Finding structure in time, Cognitive Science 14 (1990), 179–211.
- [27] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Computation 9 (1997), 1735–1780.
- [28] D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, Proceedings of the International Conference on Learning Representations, San Diego, USA, 2014.
- [29] J. Pennington, R. Socher, C.D. Manning, GloVe: global vectors for word representation, Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Doha, Qatar, 2014, pp. 1532–1543.
- [30] Y. Zhang, B. Xu, T. Zhao, Convolutional multi-head self-attention on memory for aspect sentiment classification, IEEE/CAA Journal of Automatica Sinica, 2020, pp. 1038–1044.
- [31] A. Kumar, V.T. Narapareddy, V. Aditya Srikanth, L.B.M. Neti, A. Malapati, Aspect-based sentiment classification using interactive gated convolutional network, IEEE Access 8 (2020), 22445–22453.