

METHODOLOGY

Open Access



# An aspect sentiment analysis model with Aspect Gated Convolution and Dual-Feature Filtering layers

Hongfang Gong<sup>1</sup> and Siyu Zhang<sup>1\*</sup>

\*Correspondence:  
zsyyu979@163.com

<sup>1</sup> School of Mathematics and Statistics, Changsha University of Science and Technology, Chang Sha 410114, China

## Abstract

Aspect level sentiment analysis is a basic task to determine the sentiment bias based on the contextual information near the aspect words. Some sentences contain many confusing feature words due to incomplete structure or high complexity in sentiment prediction, which can easily lead to the problem that the model pays too much attention to unimportant features. Furthermore, the role of aspect-related sentiment features is not significant in the feature extraction process. The capsule network, due to its complex network structure, leads to capsule detachment when too much information is found in the dataset. Too much computational resources are easily consumed during dynamic routing, which in turn affects the model's judgement of the polarity of aspect-related sentiment. To address these issues, we propose a capsule network (AGCDFF-Caps) with aspect-gated convolution and Dual-Feature Filtering layers for aspect-level sentiment analysis. A pre-trained BERT model is used to generate a serialised representation of the text, and the semantic representation in the contextual text is enhanced by the self-attention mechanism and bi-GRU. An aspect-gating based convolutional neural network is constructed to selectively extract contextual sentiment information about aspects and discard irrelevant information. A capsule network is then used to learn the multispatial semantic features of the aspect and the text. In particular, we incorporate a Dual-Feature Filtering network structure into the capsule network structure to strengthen the interaction between the particular aspect and the context from global and local perspectives, filtering the redundant information in local semantics and global semantics. The opinion feature representations that can more accurately express the emotional tendency of the aspect are obtained. Experimental results on SemEval2014 and Twitter datasets show that the proposed hybrid network structure has superior classification performance compared to 12 advanced baseline methods.

**Keywords:** Aspect-level sentiment analysis, Feature filtering, Aspect gating, Capsule networks

## Introduction

With the development of the Internet, the amount of data generated by users on online e-commerce platforms or online social media is increasing daily. Accordingly, mining relevant opinion information from texts with subjectivity has become important. For

example, mining users' opinions about goods from reviews on e-commerce platforms can help improve goods and services. Sentiment analysis refers to the process of analysing, processing and judging subjective text using text mining and NLP-related technologies. The traditional task of sentiment analysis can only judge the sentiment tendency of the whole sentence, whereas the task of Aspect-Based Sentiment Analysis (ABSA) [1], which is able to capture fine-grained sentiment tendencies, can better help merchants to accurately grasp the problems that exist. For example, the sentence 'Great noodles but the service is dreadful' has two aspect words, namely, the sentiment of the noodles is positive whereas the service is negative.

ABSA includes two main tasks: aspect extraction and aspect sentiment classification (ASC). We focus on the ASC task.

For the ASC task, most early feature-extraction methods rely on manual and traditional models [2, 3]. With advances in computing, deep-learning models also play good roles in NLP tasks [4, 5].

Most existing methods improve the learning performance of sentiment categorisation by constructing effective neural network structures to automatically acquire the feature information of target aspects and texts and obtain aspect-related vector representations [6–9]. Convolutional neural networks (CNNs) have excellent local feature-extraction capability and are unaffected by long-range dependency [10]. Thus, they are extensively used in ASC tasks [11–13]. Kuppusamy et al. [14] learned advanced text features by combining CNN and LSTM in a tandem manner to reduce the loss of key information. To enable CNN output vectors to aggregate more effective features, Xue et al. [15] developed a novel Tanh-ReLu unit gating mechanism for the selective output of sentiment features. Fan et al. [16] explicitly modelled word information in sentences and stored contextual information in a convolutional window to produce memory effects. Phan et al. [17] used an attentional CNN embedded with location information for aspect sentiment analysis. Lin et al. [18] used a hybrid multiple-embedding approach for aspect-level sentiment analysis on a CNN architecture.

In 2011, Hinton et al. [19] introduced capsule structures comprising multiple neurons to solve the problem of limited CNN representation, where each capsule structure represents a different attribute of the same entity. In 2017, Sabour et al. [20] proposed capsule networks, which differ from traditional CNNs in that the neurons in the capsule network are a whole and various important information are contained about the state of the feature information, such as position, angle, and orientation in the image. Conversely, traditional CNNs are insensitive to such information. Capsule networks have also achieved good results in text modelling [21, 22]. Zhao et al. [23] first applied capsule networks to a text-categorisation task. They demonstrated that capsule networks can enhance the spatial features of text while maintaining the flexibility of the textual representation. Geng et al. [24] proposed to solve the small sample text-classification problem with dynamic memory induction networks. The outcome is increased flexibility of small sample learning by using dynamic routing to better adapt to the text set. In the field of relational extraction, Zhang et al. [25] added entity and word-attention mechanisms to the dynamic routing of capsule networks to focus on discrete features. She et al. [26] proposed an interactive multi-head self-attention capsule network model (IMHSACap) to optimise the routing algorithm and activation function of the capsule

network. As opposed to focusing on optimizing the internal mechanisms of capsule networks to improve their overall performance, Wu et al. [27] focused on how to use capsule networks to better represent textual features. They reconstructed the character-level text features by using a capsule vector-based representation, which can more accurately reflect the hierarchical structure and spatial dependency among characters. Qian et al. [21] designed an interactive capsule network for the implicit sentiment analysis of text for the specific implicit sentiment expression task. Capsule networks are able to learn hierarchical features and semantic relationships in text more effectively through their unique capsule-layer structure. Its routing mechanism allows the network to dynamically adjust the connection weights between different capsules according to the characteristics of the input data. This mechanism allows the network to adaptively learn the spatial relationships among different features, thereby improving the accuracy and robustness of sentiment classification.

Despite the positive experimental results of early research in text feature extraction, some problems remain. On the one hand, CNN constructs a convolutional feature extractor to extract local features from the window, which tends to make the features redundant and inefficient in the fully connected layer. CNN also has a poor understanding of text. For example, in the ironic sentence 'His cartoons have sarcastic humour, but I like them', CNN incorrectly predicts that this sentence is a positive emotion owing to the proximity of the aspect word 'cartoon' to the opinion word 'sarcastic'. On the other hand, capsule networks achieve good results in image recognition and text modelling, but when the capsule network attempts to transmit the underlying information to higher level capsules, it may 'lose' or 'confuse' some important emotional details, especially when multiple entities or targets are involved. The emotional polarity of different entities or targets may overlap, leading to confusion in the transmission process. This can lead to confusion in the transmission of information. In dynamic routing, consuming too many computing resources is easy. In turn, the model's judgment of the polarity of aspect emotions is affected. Therefore, the present study proposes a new network model, AGCDFF-Caps, that fuses aspect gated convolution with a Dual-Feature Filtering (DFF) layer to incorporate aspect gating into the CNN before entering the primary capsule. Furthermore, to improve the effectiveness of the transfer from the junior into the senior capsule in the capsule network, a DFF network is proposed. This network exploits the interactive attention to obtain the global sequence features, and then enhances the aspect affective feature representations by using the local filtering operation and the global filtering operation. The outcome is reduced interference of redundant information.

The main contributions of this work are as follows:

- To selectively extract the sentiment information for a given aspect using aspect-related text features, we designed an aspect-gated convolution network to control aspect-related information. Aspect features are used to control textual semantic information, focusing on feature information related to aspect words.
- We incorporate DFF network in the process of capsule network passing from low-level capsule to high-level capsule, utilise richer semantic representations to deal with multilevel features, and strengthen the connection between aspect words and text from two levels of local and global features. We also retain the representations

related to the aspectual opinions through global filtering and local filtering to reduce the interference of redundant information, thereby enhancing the aspectual affective feature representations.

- The results of the experiments on the three publicly available datasets used for the experiments are as follows: Restaurant and Laptop data from SemEval 2014 Task4 and the Twitter dataset from ACL14 perform better than the baseline model, demonstrating the effectiveness of the AGCDDFF-Caps model in aspectual sentiment-analysis tasks.

The rest of the paper is organised as follows. Sect. "[Related work](#)" briefly describes the work related to the sentiment-analysis task. Sect. "[Our approach](#)" describes the components of the AGCDDFF-Caps model. Sect. "[Experiments](#)" verifies the validity of the AGCDDFF-Caps model and its modules against existing baseline models. Sect. "[Experimental results and analysis](#)" summarises the work.

## **Related work**

### **Aspect-level sentiment analysis**

The purpose of the ABSC is to determine the affective tendency of entity aspect words in context. It has achieved significant results in previous sentiment-classification studies [28, 29]. Early ABSC tasks are computationally complex and labour intensive with low accuracy [30, 31]. In recent years, neural networks have achieved good results in ABSC owing to their self-learning capability and fast computational speed [30–34]. Wang et al. [35] introduced aspect words containing LSTM context information while computing attention weights. Xu et al. [36] used Transformer to extract hidden features of aspect and text after GloVe encoding. Interactive learning by two-channel global and local attention extracts interaction information of different granularity, embeds location coding into the attention mechanism, and assigns corresponding weights to the extracted feature information. Zhou et al. [37] modelled location information in different ways and improved the performance of the model by transferring hierarchical knowledge information. Xu et al. [38] proposed a new concept of dependency clustering and devised contextual concerns and dependency clustering concerns, focusing on the more critical clustering. Phan et al. [17] used a graphical CNN to simultaneously investigate contextual information, semantic relationship information, and sentiment knowledge between words or phrases. Xiao et al. [39] used RGAT to encode the structure of the tree and also introduced a dense graph convolutional network to consider the distances between aspect and related words. They aimed to better establish and strengthen the link between aspect and sentiment word.

### **Gated networks**

The emergence of a gating mechanism [40] is good for improving the selection of important information in a sentence. Parveen et al. [41] proposed a bidirectional recurrent neural network (BRNN). This model can use old and current information while introducing bidirectional gated recurrent units (bi-GRUs) to replace the hidden layer of the BRNN with a single GRU storage unit and subsequently control the flow of information. Kumar et al. [42] selectively learned features through a self-attention and gating

mechanism, where features contain less information flow and gating focuses on interactions. Liu et al. [43] designed a filter gate between the bi-GRU and the attention layer to control the flow of information from layer to layer and thus attenuate the influence of useless information. Kumar et al. [44] used a bidirectional interactive gated convolutional network to learn the relationships between aspect and context. Ran et al. [45] constructed a hierarchical gate memory network that selectively outputs aspect-related contextual representations of a given sentence. Experiments show that it outperforms models based on attention mechanisms. However, the above studies are more concerned with gated filtering of text, and gating mechanisms have also achieved good results in aspect-level sentiment-classification tasks. Lu et al. [46] introduced a gating mechanism in Bi-LSTM to guide the encoding of aspect-related sentiment information, which is then used to capture the long-range dependencies between words via GCN. Han et al. [47] used a self-attention mechanism to obtain aspect features and then dynamically fused aspect-related information using an information gate based on a gating mechanism. Kamil et al. [48] used a gated recurrent unit (GRU) in combination with multiple feature-extraction methods for aspect-level sentiment analysis.

### Capsule networks

The emergence of capsule networks is a good solution to the problem of insufficient learning of spatial information by models [49–51]. Wang et al. [52] used capsule neural networks to identify spatial and semantic relationships between features to extract implicit features. Su et al. [50] constructed auxiliary sentences by using XLNetCN to model sequence–aspect relations. Then, they extracted the local and spatial hierarchical relations of text sequences through a dynamic routing algorithm in capsule networks and generated their local feature representations. They are merged with the global aspect-related representations for downstream classification using softmax classifiers. Wang et al. [53] applied aspect-level sentiment features to capsule networks, a model that exploits the correlation between aspects and sentiments and performs ABSC tasks in parallel through shared components. Zhang et al. [54] guided capsule networks by dynamically adjusting the information transfer of prior knowledge to improve text representations. Yang et al. [22] applied dynamic routing to a general neural network to capture sentiment features between aspect and opinion words. Lin et al. [55] proposed an elementary discourse unit Capsule network (EDU-Capsule) for ABSC. However, the high performance of the capsule network cannot be separated from its dynamic routing. When the data set becomes huge, the capsule easily falls off the situation and consumes too much computer resources. In the process of capsule-feature transfer, if several similar types of objects exist in the text and they are very close to one another, the advanced capsule may not be able to accurately distinguish the correct features. This situation is called a ‘congestion phenomenon’, and the performance of the task requires further improvement.

The high performance of the capsule network cannot be separated from its dynamic routing. When the data set becomes huge, the capsule tends to fall down and consume too much computer resources, and its task performance needs to be further improved. Based on the above study, this paper combines the improved capsule network with the gating mechanism while incorporating aspect capsules to learn aspect-overlap

information. The flow of aspect related information is controlled by aspect-gating convolution to enhance aspect features. The gating network with filtering properties is used to achieve feature enhancement of global and local features, reducing the interference of redundant information during model learning.

### Our approach

Given a sentence  $S = w_1, w_2, \dots, w_n$  and an aspect item  $T = w_{a_1}, w_{a_2}, \dots, w_{a_m}$ , our task is to predict the affective tendency of aspect item  $T$ , i.e. positive, negative, or neutral. The opinion information associated with the aspect item in sentence  $S$  is learning. We then describe the model as a whole. In this part, we first provide an overview of the whole model and then describe the different parts of the model in detail.

### Framework

The proposed AGCDDFF-Caps model comprises five main modules: (1) an embedding layer, which obtains the vector representation of context and aspect, respectively, through the BERT pre-training model; (2) a semantic extraction layer, in which the combination of self-attention mechanism and bi-GRU is used to capture the semantic correlations and enhance the semantic information of each word in the sentence; (3) an aspect-gating convolution layer, which applies the gating unit to control the aspect-related information; (4) a capsule network layer, where low-level capsules are passed to high-level ones by iteratively renewing the coupling coefficients and aggregating the capsule information to output the classification probability; and (5) a dual-feature filtering layer, which performs a distillation-like operation on global and local sequences to retain the aspect-related sentiment features. An overview of the AGCDDFF capsules is shown in Fig. 1.

### Embedding layer

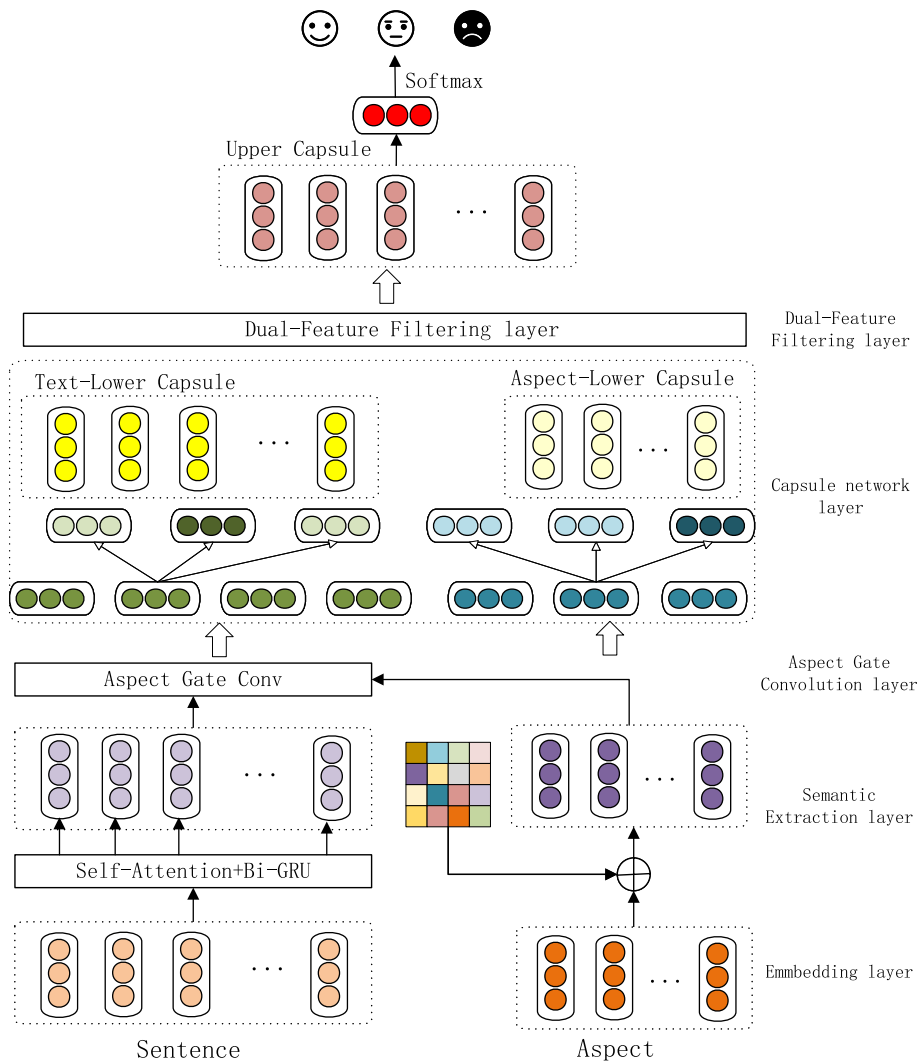
BERT as a language model is represented by a bidirectional encoder of the Transformer with the classification symbol [CLS] at the start position and the separator character [SEP] in the break interval, e.g. '[CLS]+ Sentence1 + [SEP] + Sentence2'. BERT can capture vector features of sentences and aspects. We feed the sentence  $S$  and the corresponding aspect  $T$  into BERT to obtain  $X^S = \{x_{s_1}, x_{s_2}, \dots, x_{s_m}\}$  and  $X^T = \{x_{a_1}, x_{a_2}, \dots, x_{a_n}\}$ , respectively, where  $x_i \in R^e$  is an  $e$ -dimensional word vector.

$$X_S = BERT(S) \quad (1)$$

$$X_T = BERT(T) \quad (2)$$

### Semantic extraction layer

The role of the semantic extraction layer is to attend to and extract information about the global context and important semantic information related to aspects. The GRU controls inputs and outputs through reset gates and update gates and subsequently uses gating mechanisms to allow the model to learn the dependency state at the current moment. The self-attention mechanism has a large sensory field and can capture the semantic relationship between any two elements in a sequence to enable contextual



**Fig. 1** The overview of AGCDFF-Caps

interaction and integration. To capture long-range contextual information, we first use a bi-GRU to obtain contextual and aspect features of the text. We then augment the text semantics by using the self-attention mechanism and multiply the obtained attention score matrix with the aspect features. Accordingly, we obtain feature information that is closely related to the aspects.

$$r_t = \sigma(U_r[h_{t-1}; x_t^S]) \tag{3}$$

$$z_t = \sigma(U_z[h_{t-1}; x_t^S]) \tag{4}$$

$$h'_t = \tanh(U_h[r_t \times h_{t-1}; x_t^S]) \tag{5}$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times h'_t \tag{6}$$

$$H^S = \vec{h}^S \oplus \overleftarrow{h}^S ; H^T = \vec{h}^T \oplus \overleftarrow{h}^T \quad (7)$$

$$Attention = softmax\left(\frac{QW^Q \times (KW^K)^T}{\sqrt{d}}\right) \quad (8)$$

$$H^t = Attention \cdot H^T \quad (9)$$

where Eqs. 3–6 denote the calculation formulas of the reset gate and update gate at the moment  $t$ ;  $[x; y]$  denotes the splice of the two vectors;  $\sigma$  denote activation function;  $U_t$ ,  $U_z$ , and  $U_h$  denote the weight matrices;  $r_t$  and  $z_t$  are the outputs of the update gate and reset gate, respectively;  $h_t$  and  $h_{t-1}$  denote the hidden layer outputs of the GRU at the moments  $t$  and  $t-1$ , respectively;  $\vec{h}$  and  $\overleftarrow{h}$  are the final hidden layer representations of the GRU learning from the back-end sequence and GRU learning from the back-end sequence, respectively;  $\oplus$  denotes the splice of two vectors; Attention is the attention score matrix;  $H^t$  is the aspect feature of the final output; and  $H^S$  is the context feature of the final output.

$$H^t = \{h_1^t, h_2^t, \dots, h_n^t\} \quad (10)$$

$$H^S = \{h_1^S, h_2^S, \dots, h_m^S\} \quad (11)$$

### Aspect-gated convolution layer

In this layer, we can use aspect-related text features to selectively extract the sentiment information for a given aspect. With the gating mechanism, the contextual sentiment information related to the aspect is retained while the irrelevant information is discarded. In the example ‘This restaurant is good but the delivery is slow’, if the gate controls the information flow of the first aspect ‘restaurant’, it selectively ignores the sentiment information conveyed by the next sentence and outputs the sentiment information only of the previous sentence.

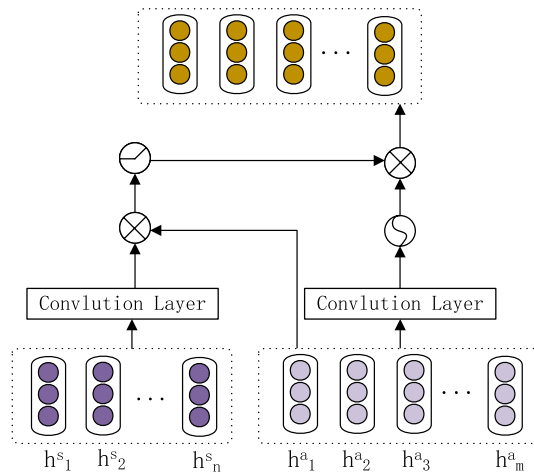
Figure 2 illustrates the framework of our aspect-gated convolutional layer. The gated activation units used for aspect information flow are connected to the convolution neural network. The context features  $h_i^S$  are convolved to obtain locally enhanced text features  $H_{conv}^a$  and then dot multiplied by  $h_j^a$  to obtain aspect-enhanced features with reassigned weights from the aspect features. The ReLU function is used to control the aspect-independent noise information to obtain the feature  $a_i$ . The input aspect features  $h_j^a$  are also convolved to obtain locally enhanced aspect features and then mapped to the value within the range of -1 to 1 by using the Tanh function. It helps the model to better learn and represent aspect-related emotional features. The specific calculation process is as follows:

$$H_{conv}^a = Conv(h_i^S \otimes W_1) \quad (12)$$

$$a_i = ReLU(H_{conv}^a + W_2 h_j^a + b_1) \quad (13)$$

$$b_i = Tanh(Conv(h_j^S \otimes W_3) + b_2) \quad (14)$$





**Fig. 2** Overview of the Aspect Gated Convolution Network

$$h_i^{Sg} = a_i \times b_i \tag{15}$$

where  $W_1, W_2, W_3 \in \mathbb{R}^{h \times d}$  is the weight matrix,  $a_i$  denotes aspect features, and  $b_i$  denotes affective features.

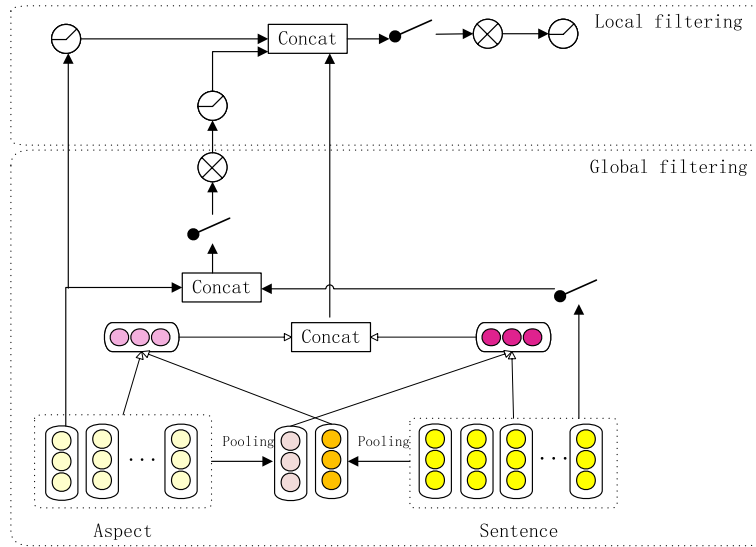
**Dual-feature filtering layer**

In capsule networks, the contextual features of multiple capsules obtained from primary capsules belong to the coarse-grained text representation. On one hand, the primary capsule provides a rich feature representation for the advanced capsule by encoding spatial and directional information in the capsule network, which helps improve model performance. On the other hand, complex feature relationships tend to mislead the model into focusing on useless features, introducing additional noise and losing important information. Therefore, DFF is used to construct a feature-filtering channel to deeply extract key information related to aspects correct and weaken the features with low correlation between aspect words and text. The results of the two channels are then stitched to obtain the accurate expression of aspect sentiment.

The specific framework of DFF is shown in Fig. 3. The entire network is constructed based on the gating mechanism, which is divided into two parts: global filtering and local filtering. In global filtering, the Sigmoid activation function is applied to the spliced features of text feature  $h_s$  and aspect feature  $h_t$  and to the single-text feature  $h_s$ . The function can map values between 0 and 1, reflecting the degree of correlation between global and aspectual features. Then,  $h_{ss}$  is dot multiplied to obtain the global features  $h_{gs}$  with the weights reassigned by the text features, and the spurious information in them is filtered out using the ReLU function. The specific calculation is as follows:

$$h_{ss} = \sigma(W_{ss} \cdot h_s + b_{ss}) \tag{16}$$

$$h_{gs} = \sigma(W_{gs} \cdot [h_{ss} \oplus h_t] + b_{gs}) * h_{ss} \tag{17}$$



**Fig. 3** Overview of the Dual-Feature Filtering network

$$h'_{gs} = \text{ReLU}(W'_{gs}h_{gs} + b'_{gs}) \tag{18}$$

where  $h_{ss}$  is the text feature obtained by applying the activation function;  $h_{gs}$  is the initial global feature obtained by activation after splicing;  $h'_{gs}$  is the global feature vector obtained after global filtering; and  $W_{ss}$ ,  $W_{gs}$ ,  $W'_{gs}$  and  $b_{ss}$ ,  $b_{gs}$ ,  $b'_{gs}$  are the learnable weight matrix and the deviation vector.

In local filtering, to strengthen the connection between aspect words and context, an interactive attention mechanism is introduced. It integrates the aspect terms and the corresponding textual information while extracting deeper information related to the aspect. Afterwards, it performs interactive learning modelling on both separately. The final learned feature representations are spliced and aggregated to obtain the strengthened aspect feature vectors  $h_{ts}$ . Then, the activated aspect feature  $h'_t$  is subjected to a filtering operation, and the weights of the local feature  $h'_{gs}$  are assigned by the global feature  $h'_t$ . Finally the aspect-specific sentiment feature  $h_a$  is obtained after ReLU activation. The computational process is as follows:

$$h_{as} = \text{InterAtt}(h_a, h_s) \tag{19}$$

$$h'_{as} = \sigma(W'_{as} \cdot h_{as} + b'_{as}) \tag{20}$$

$$h'_a = \text{ReLU}(W'_a h_a + b'_a) \tag{21}$$

$$h_{la} = \sigma(W_{la}[h'_{gs} \oplus h'_{as} \oplus h'_a] + b_{la}) * h'_{gs} \tag{22}$$

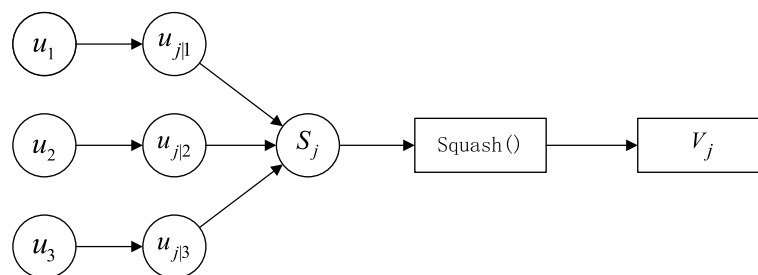
$$H_a = \text{ReLU}(W_d(h_{la} + W_{ad}h'_a) + b_d) \tag{23}$$

where  $h'_{as}$  is the enhanced aspect feature obtained by applying the activation function;  $h_{la}$  is the feature vector obtained by interacting the globally filtered feature vector with the new aspect features;  $h_a$  is the final output vector obtained after two filters; and  $W'_t$ ,  $W'_{ts}$ ,  $W_{la}$ ,  $W_d$ ,  $W_{ad}$  and  $b'_t$ ,  $b'_s$ ,  $b_{la}$ ,  $b_d$  are the learnable weight matrix and the deviation vector, respectively.

### Capsule network layer

A capsule network is a collection of neurons whose vectors represent the probability of occurrence of certain entities or attributes. The primary capsule passes through the transformation matrix to the senior capsule. Higher-level capsules are activated when multiple predictions match. Capsule networks have a more powerful feature representation than CNNs. The capsules in the capsule network can encode and encapsulate the attributes of different entities in the image, such as position, orientation and angle. This makes the capsule feature representation more specific and richer. Conversely, feature representations in CNNs are often based on the pixel level, conferring difficult in directly representing the specific attributes of entities in an image. In sentiment analysis, capsule networks can better understand the semantic relationships and hierarchical structures in text and can thus more accurately identify the emotional tendencies in text. For example, in social-media text, capsule networks can capture the emotional associations between different sentences or paragraphs, so it can more accurately determine the emotional tendency of the entire text.

In general, a capsule network comprises a primary capsule and a senior capsule. The primary capsule determines the output of the senior capsule, which primarily comprises a dynamic routing that iteratively updates the weight matrix, and the final output capsule vector comprises all senior capsules. Owing to its special dynamic routing update process, the capsule network can extract deeper text features. Its network structure is shown in Fig. 4.



**Fig. 4** Capsule network structure

**Algorithm 1** Algorithm 1 Dynamic routing algorithm

---

```

1: procedure ROUTING( $\hat{u}_{j|i}, r, l$ )
2:   for all capsule  $i$  in layer  $l$  and capsule  $j$  in layer  $(l + 1)$ :  $b_{ij} \leftarrow 0$ .
3:   for  $r$  iterations do
4:     for all capsule  $i$  in layer  $l$ :  $c_i \leftarrow \text{soft max}(b_i)$ 
5:     for all capsule  $j$  in layer  $(l + 1)$ :  $s_j \leftarrow \sum_i c_{ij} * \hat{u}_{j|i}$ 
6:     for all capsule  $j$  in layer  $(l + 1)$ :  $v_j \leftarrow \text{squash}(s_j)$ 
7:     for all capsule  $i$  in layer  $l$  and capsule  $j$  in layer  $(l + 1)$ :  $b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i} \cdot v_j$ 
   return  $v_j$ 

```

---

Our study uses a capsule network based on a dynamic routing algorithm, where an affine transformation is performed in the primary capsule layer, and the features that have passed through the convolutional layer are initialised and multiplied by a weight matrix. This matrix is then mapped into a vector space with the same number of neurons as the capsule layer:

$$\hat{u}_{j|i} = W_{ij}u_i \quad (24)$$

where  $\hat{u}_{j|i}$  is the prediction of a single low-level capsule to a single high-level one.

We calculate the total input  $S_j$  to each high-level capsule from the low-level capsule:

$$S_j = \sum_i c_{ij} * \hat{u}_{j|i} \quad (25)$$

where  $c_{ij}$  is the coupling coefficient, indicating the degree of polymerisation of the low-level capsule  $u_i$  with the high-level one  $u_j$ . The sum of the coupling coefficients is usually given as 1; and  $c_{ij}$  is renewed iteratively by dynamic routing, calculated as follows:

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (26)$$

$$b_{ij} = b_{ij} + \hat{u}_{j|i} \cdot v_j \quad (27)$$

where  $c_{ij}$  is obtained by softmax, which denotes the weight of the weighted sum; and  $\hat{u}_{j|i} \cdot v_j$  denotes the similarity between the low level capsule  $i$  and the high-level one  $j$ .

Corresponding with each high-level capsule  $j$ , the nonlinear squeeze function *squash* is applied to compress the length of the entire input vector to obtain the output feature vector  $v_j$ :

$$v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} * \frac{s_j}{\|s_j\|} \quad (28)$$

For each low-level capsule  $i$  and high-level one  $j$ , the a priori coefficients  $b_{ij}$  are updated based on the dot product between their output vectors  $v_j$  and eigenvectors  $\hat{u}_{j|i}$ .  $c_{ij}$  is then updated by  $b_{ij}$ , i.e., the coupling strength between the low-level capsule and the high-level one is adjusted based on the feedback from the latter.

The feature vectors are transmitted into the fully connected layer to output the predicted sentiment.

$$y = \text{Soft max}(W_a v + b_a) \quad (29)$$

where  $W_a$  and  $b_a$  are the learnable weight matrix and bias vector, respectively.

The model parameters are then optimised by minimising the cross-entropy loss between the true and predicted values, whereas L2 regularisation is used to prevent model over-fitting.

$$L(\theta) = - \sum_{i=1}^N y_i \lg(\hat{y}_i) + \lambda \|\Theta\|^2 \quad (30)$$

where  $N$  is the total number of samples,  $y_i$  is the true ASC,  $\hat{y}_i$  is the predicted ASC,  $\lambda$  is the L2 regularisation coefficient, and  $\Theta$  denotes all learnable parameters of the model.

## Experiments

### Experimental dataset

To evaluate the validity of our proposed model, experiments are conducted on three public social-platform datasets, namely, the ACL2014 Twitter and the Laptop, Restaurant datasets from SemEval2014 Task 4. Each sentence in these datasets contains multiple aspects corresponding with sentiment polarity. Table 1 shows the situation for each specific data.

### Experimental-parameter settings

In the experiments, accuracy and F1 value are used as indices for model evaluation. Accuracy is the percentage of correct predictions for all sample categories. The F1 value better reflects whether the performance of the classifier is robust or not.

The model is built in the Pytorch framework environment. The Adam optimiser is used with the learning rate set to  $5e-5$  and the L2 regularisation coefficient set to  $1e-5$ . The word vector dimension is 768, the hidden layer state vector dimension is 300 and the batch\_size is 32.

In this paper, we use the Bert model to pre-train the sequential text, the hidden layer dimension is 768, and the maximum sentence length is 120. The hidden layer state vector dimension of bi-GRU is 384. The Train\_Batch\_size of the dataset Restaurant and Laptop is 32, and the Eval\_batch\_size is 16. The dataset Twitter has a Train\_Batch\_size of 32, and the Eval\_batch\_size is 16. Twitter has a Train\_Batch\_size of 32, and an Eval\_Batch\_size of 32. The model is built using the Adam optimiser with

**Table 1** Statistical data for the three data sets

Dataset	Restaurant		Laptop		Twitter	
	Train	Test	Train	Test	Train	Test
Positive	2164	728	994	341	1561	173
Neutral	807	196	870	128	3127	346
Negative	637	196	464	169	1560	173

a learning rate of  $5e-5$  and a dropout of 0.3. The model is built using the pytorch environment. The details of the setup are shown in Table 2. The optimal parameters based on the best performance are kept on the development set, and the optimal model is used to evaluate the test set.

### Baseline methods

To measure the feature-extraction performance of the models, we compare them with previous state-of-the-art neural networks. Experimental results are taken from literature corresponding with each model.

TD-LSTM [56] learns aspect-related representations by feeding target information to an extended LSTM.

PBAN [57] learns relationships between aspect words and text using bidirectional attention.

RAM [6] uses a multiple-attention mechanism to capture distant emotional information and improve the ability to learn multi-information representations.

AEN [58] enhances semantic relations between context and aspect using a coder based on multiple attention mechanisms.

DGEDT [59] learns representations from semantic graphs by iteratively interacting transformers.

TNet [60] extracts features using CNNs and bidirectional RNNs and proposes transformations for the target aspect representation to better represent aspect information.

XLNet-CNN-GRU [61] generates dynamic stacked word vectors via Glove and XLNet and then feeds the stacked word vectors into a two-channel GRU and CNN with an attention mechanism to effectively solve the phenomenon of word polysemy.

ABWE [62] adopts an aspect-based word embedding method to filter the aspect sentiment word, It uses the bi-GRU method to obtain global context information, which is merged with aspect and opinion information to obtain feature representations about words.

KGCNCapsAN [54] introduces an attention mechanism to improve capsule networks while introducing a priori knowledge to guide the routing of capsule attention networks.

**Table 2** Experimental parameter setting

	Restaurant	Laotop	Twitter
Hidden Layer Dimension	768	768	768
Maximum sentence length	120	120	120
Train_Batch_size	16	16	32
Eval_batch_size	32	32	32
Learning rate	$5e-5$	$5e-5$	$5e-5$
Dropout	0.3	0.3	0.3
Epoch	30	30	30

TransCap [63] encapsulates document-level knowledge in aspect-sentiment capsules. It also adaptively combines semantic capsules and categorisation capsules through migration learning to improve the learning performance of the model.

AOIARN [22] uses interactive learning to learn the correlation between specific aspects and emotion words using BiLSTM. It embeds dynamic routing in neural networks to improve emotion-information extraction.

EDU-Capsule [55] utilises capsule networks for aspect-level sentiment analysis in terms of grammatical units of sentence clauses.

## Experimental results and analysis

### Effectiveness of AGCDFF-caps

The experimental results of the AGCDFF-Caps model and the previous baseline approach on the dataset are shown in Table 3. TD-LSTM by modelling the context of aspect words is aspect information is more focused, but the obtained sentiment is biased towards the whole. Conversely, AEN, PBAN, RAM, and DGEDT fine-grained analysis methods based on attention capture the semantic information between aspect and context with higher performance. However, the learning of the attention mechanism here is limited to the self-learning of a single text without interactive modelling with aspectual words. The lack of attention to the contextual positional information of the aspectual words may not allow for the effective fusion of the information of the aspectual words in the textual features. TNet, XLNet-CNN-GRU, and ABWE, represented by embedding bi-GRU based on gating mechanism into neural networks, can improve the training efficiency of the model when dealing with long sequential data. They capture the feature information in sequential data more comprehensively by considering past and future information. These models lack a self-attention mechanism compared with AGCDFF-Caps and are thus unable to dynamically adjust the

**Table 3** Experimental results for the restaurant, laptop, and Twitter datasets

Method		Restaurant		Laptop		Twitter	
		Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
Deep learning network	TD-LSTM(2015)	75.60	64.50	68.10	63.90	66.60	64.00
	RAM(2017)	78.48	77.32	72.08	68.43	70.09	66.48
	PBAN(2018)	78.62	67.45	71.98	66.91	-	-
	TNet(2018)	80.70	71.30	76.50	71.80	75.00	73.60
	AEN(2019)	80.99	72.14	73.51	69.04	72.38	69.81
	DGEDT(2020)	83.90	75.10	76.80	72.30	74.80	73.40
	XLNet-CNN-GRU(2023)	82.67	75.32	75.87	73.41	73.42	71.59
	ABWE(2023)	80.25	78.44	71.42	69.53	-	-
Capsule network	TransCap(2019)	79.55	71.41	73.51	69.81	-	-
	KGCNCap-sAN(2020)	82.05	74.04	76.96	72.89	74.13	72.52
	AOIARN(2022)	83.40	75.05	78.90	75.10	75.00	73.80
	EDU-Capsule(2023)	85.70	-	78.50	-	-	-
Ours	AGCDFF-Caps	<b>86.16</b>	<b>79.36</b>	<b>80.76</b>	<b>76.17</b>	<b>75</b>	<b>73.81</b>

The best results are shown in bold

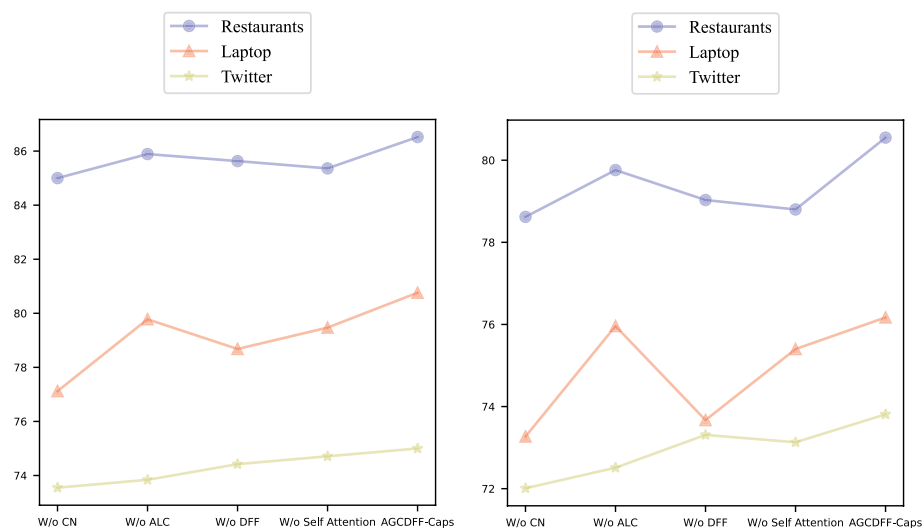
degree of attention to other positional information in the sequence directly based on the information at the current position. Consequently, the model’s effective integration of global contextual information is limited. Furthermore, XLNet-CNN-GRU has a more significant performance improvement than ABWE and TNet owing to the fact that XLNet-CNN-GRU introduces the ability of CNNs to focus on locally important features, consider multiple semantic feature information, and strengthen the model’s ability to learn aspects of relevant sentiment information.

The baseline models KGCNCapsAN, TransCap, AOIARN, and EDU-Capsule use the dynamic routing process of the capsule network to learn to update the weight matrix and extract deep semantic information. The main difference from AGCDFF-Caps is that this paper’s model models the important sequence semantics while filtering the redundant information several times to focus on the sentiment expressions in the model’s output results that are closely related to aspects. Experimental results show better performance of this paper’s model, indicating that the classification effect after DFF on the capsule network is better than that of the model without filtering operation, and the fine aspect-sentiment features are more conducive to improving the model-learning performance than the features with redundant information. Thus, the effectiveness of the hybrid AGCDFF-Caps network is verified.

**Ablation experiment**

To assess the should of each module of AGCDFF-Caps on the classification effect, ablation experiments are performed separately for each module within the model:

- W/o Capsule Network removes the whole capsule network structure,
- W/o Aspect-Lower Capsule removes the aspect-lower capsule structure,
- W/o DFF layer removes the DFF structure, and,
- W/o Self Attention removes the self-attention mechanism.



**Fig. 5** Results of ablation experiments. The left panel shows the accuracy results, and the right panel shows the F1 value results



The results are shown in Fig. 5. Removing any component leads to decreased accuracy and MF1 value of the model.

Figure 5 shows that removing any part of the components within the model leads to decreased accuracy and F1 value of the model, indicating that each module within the model has a positive effect. However, the capsule network has the greatest impact on the model performance. Capsule networks act as a neural network structure that captures and represents hierarchical relationships among entities. In text processing, capsule networks can effectively model and integrate text features at different levels such as words, phrases, and sentences to improve the model's ability to understand deeper information in the text. By integrating capsule networks into the model, we can allow it to focus more on learning the key information in the text and ignore the noisy information that is not important for the task.

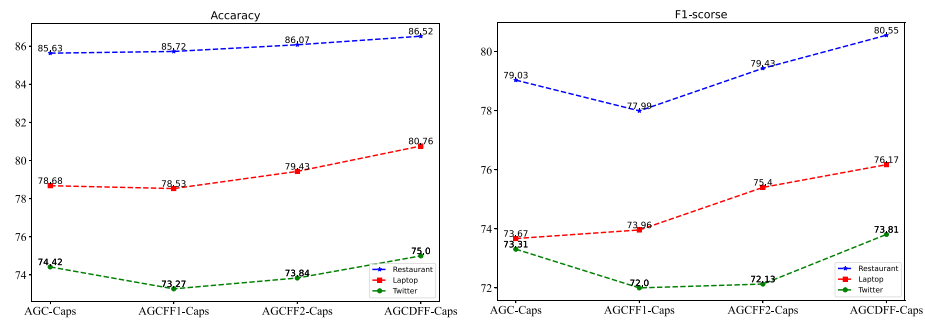
In the model W/o Aspect-Lower capsule compared with W/o capsule networks, the accuracy increases by 0.89%, 2.66%, and 0.29% respectively. This finding indicates that the aspect features also contain important information related to the sentiment opinions, which is one of the key elements to understand the sentiment tendency of aspect words. Additionally, the accuracy of the model w/o DFF layer shows different degrees of decrease, specifically 0.89%, 2.08%, and 0.58%. This result highlights the importance of the DFF structure in the model, especially the powerful ability to filter useless information. DFF can effectively filter and integrate key information related to aspectual sentiment through global feature filtering and local feature filtering to improve the accuracy of the model.

In the Laptop dataset, we observe a particularly significant change in learning performance. This change is primarily owing to the missing or incomplete sentence information in this dataset. The sentence expressions in the laptop dataset are usually more concise and lack sufficient contextual information, which relatively limits the ability of the model to filter and integrate information.

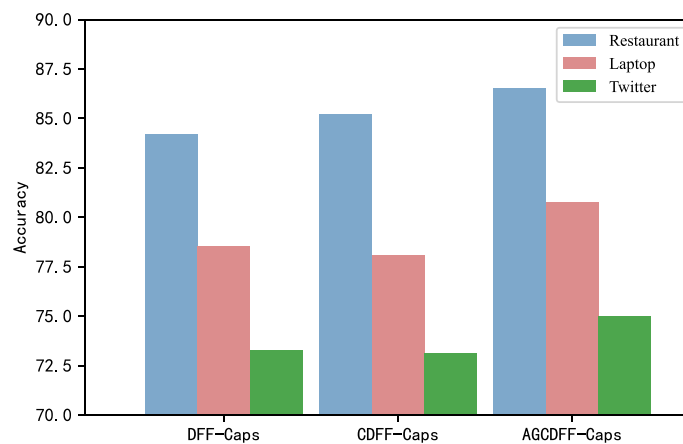
Furthermore, when we remove the self-attention mechanism from the semantic extraction phase, the performance of the model significantly decreases. This finding highlights the importance of the self-attention mechanism in capturing the semantic dependencies of each word in a sentence. The self-attention mechanism allows the model to dynamically compute the dependencies between different positions in a sequence while processing textual data and generates a weight matrix for representing the importance of each position in the sequence. This capability allows the model to more accurately understand the overall structure and semantic information of the sentence.

#### **DFF module-performance analysis**

To further investigate the effectiveness of the DFF module in blocking redundant information, we design four variants of the model: the AG-Caps (removing the entire DFF portion), the AGCFF1-Caps (retaining the local feature-filtering channel and removing the global feature-filtering channel), the AGCFF2-Caps (retaining the global feature-filtering channel and removing the local feature-filtering channels), and the AGCDFF-Caps. AGCFF1-Caps shows faster model-performance degradation than AGCDFF-Caps. This finding suggests that the global feature-filtering channel plays a more critical role in capturing sentiment polarity. Global feature filtering integrates all information in the



**Fig. 6** Comparison of accuracy and F1 value for different feature-filtering channels



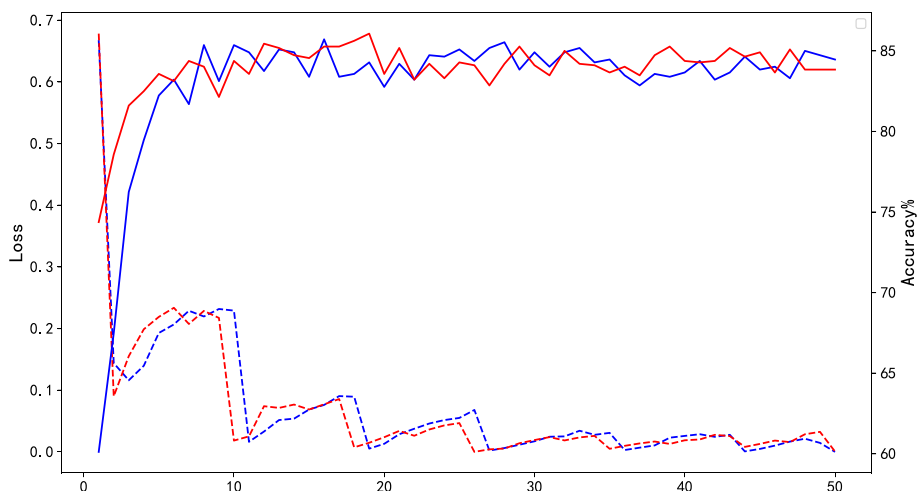
**Fig. 7** Comparison of experimental results for different gating mechanisms

sentence and the context to more accurately identify key features relevant to sentiment classification. When this channel is removed, the model’s ability to detect sentiment polarity is significantly impaired. The results of AG-Caps as a model that removes the entire DFF further validate the effectiveness of the DFF module in blocking redundant information and improving model performance.

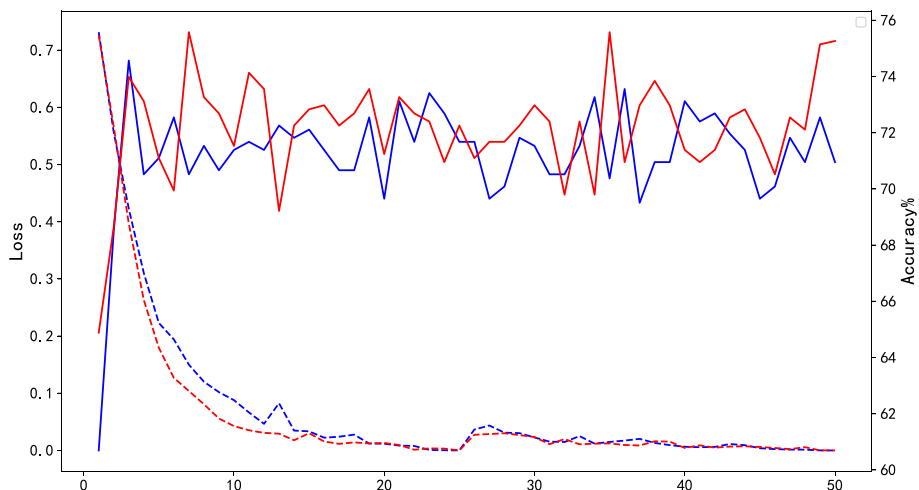
DFF plays an important role in blocking redundant information and improving the performance of sentiment classification. Global feature filtering and local feature filtering are complementary in the sentiment-classification task. The combination of the two can capture the sentiment information in a sentence more comprehensively (Fig. 6)

**AGC module-performance analysis**

The experimental results in Fig. 7 are compared with the three methods, DFF-Caps (removing the entire aspect-gated convolution), CDFE-Caps (removing the aspect-gating mechanism), and AGCDFF-Caps. Results further demonstrate the advantages of Aspect-Gated convolution in improving the performance of sentiment analysis models. Comparing DFF-Caps without CNNs to CDFE-Caps with CNNs, we find that the performance of DFF-Caps is better on the Laptop and Twitter datasets. This result reveals the limitations of CNNs in feature extraction. When the convolutional kernel is too small, it cannot establish the temporal connection well; when the convolutional kernel



**Fig. 8** Comparison of accuracy and convergence on the Restaurant dataset



**Fig. 9** Comparison of accuracy and convergence on the Twitter dataset

is too large, too many features are extracted and noise generation ensues. Compared with CDFD-Caps, AGCDFD-Caps significantly improves the model’s ability to learn high-quality aspect-feature information by applying the aspect-gating mechanism. The aspect-gating mechanism can effectively control the information flow and accurately consider the sentiment opinion information related to aspects. Thus, the model can pay more attention to task-relevant feature information, which helps it more accurately identify the key features related to sentiment classification. This mechanism enables AGCDFD-Caps to achieve higher accuracy in sentiment-analysis tasks.

**Analysis of capsule structural performances**

To investigate the effect of the DFF layer on the overall performance of the model, AGCDFD-Caps with DFF layer is experimentally compared with AG-Caps without DFF on the datasets Restaurant and Twitter. The convergence curves with the number of

training iterations and the accuracy curves are shown in Figs. 8 and 9, respectively, where the red realisation and the red dashed line correspond with the accuracy and loss values of the AGDFF-Caps model, and the blue solid line and the blue dashed line correspond with the accuracy and loss values of the AG-Caps model.

The experimental results on the two datasets show that the AGCDFF-Caps model proposed in this paper has a significant improvement in accuracy compared with the AG-Caps model without DFF. This improvement proves the effectiveness of DFF in feature extraction and further emphasises the importance of feature screening and filtering in sentiment-analysis tasks. Specifically, the DFF in the AGCDFF-Caps model is able to prioritise the filtering of features as they are passed from lower level capsules to higher level capsules. This filtering mechanism allows feature information that is aspectually relevant and strongly expresses emotion to be retained and passed on to subsequent processing. In this way, the model can use limited computational resources more efficiently and focus on processing key features with significant impact on sentiment classification.

Our experiments also show that the AGCDFF-Caps model converges faster. This finding further confirms the advantages of DFF in feature filtering. Given that DFF can preliminarily remove features that are irrelevant or noisy for the task, the model is able to converge to the optimal solution faster during training. This efficient performance makes the AGCDFF-Caps model more competitive in sentiment-analysis tasks, enabling better contextual representations faster.

## Conclusions

We propose a capsule network model incorporating aspect-gated convolution and DFF layers for aspect-level sentiment analysis. The model solves the problems of confusing other feature words during feature extraction, resulting in insignificant aspect-related sentiment features and feature overlap or redundancy during capsule network delivery, as well as providing programmatic guidance for businesses, governments, and organisations, among others.

The model can utilise the self-attention mechanism and bi-GRU for deep semantic feature extraction. It also strengthens the aspect feature representation by aspect-gated convolution, and establishes the text capsule network and the aspect capsule network, respectively. In the capsule network structure, to enhance the effectiveness of the transfer process from low-level capsules to high-level ones in the capsule network, we incorporate a DFF layer, a design that can sequentially perform feature filtering on global sequence features and local sequence features. Accordingly, the impact of unimportant features on the learning aspectual sentiment of the model is effectively reduced. Experimental results on three publicly available datasets show that the model proposed in this paper performs better in terms of performance, and the aspect-gated convolution successfully enhances the aspect features that are closely related to the task of sentiment analysis. Thus, the model more accurately captures the sentiment tendencies in the text. The introduction of the DFF layer further improves the feature-selection ability of the model, making the model more robust in dealing with complex sentiment-analysis tasks.

Although the model in this paper shows superior performance and good interpretability in aspect-level sentiment analysis tasks, there are still some shortcomings that need to be further explored and improved in our future research. Firstly, the model may

not perform well for some specific types of sentiment-analysis tasks. For example, when dealing with fine-grained sentiment-analysis tasks, the model may not be able to accurately capture subtle sentiment differences in the text. To further improve the model performance, we can consider introducing more contextual information or adopting a more complex network structure to capture subtle sentiment changes in the text. Secondly, the generalisation ability of the model is a key concern. The model achieves good performance on the three public datasets, but it may encounter inconsistencies with the distribution of training data in practical applications. To improve the generalisation ability of the model, we can use data-enhancement techniques or apply unsupervised learning methods to pre-train the model so that it can better adapt to different application scenarios.

#### Acknowledgements

We are very grateful to the anonymous reviewers and the editor-in-chief for their valuable comments, which greatly help improve the quality of this paper. This work was supported in part by the National Natural Science Foundation of China under Grant 61972055 and in part by the Natural Science Foundation of Hunan Province of China under Grant 2021JJ30734.

#### Author contributions

S.Y. provided methodological theory, performed data analysis, data visualisation, prepared Figs. 1–9 and wrote the main manuscript. G.F. provided methodological guidance, process supervision and wrote the main text of the manuscript. All authors revised the manuscript and validated the methodology.

#### Availability of data and materials

No datasets were generated or analysed during the current study.

#### Declarations

##### Competing interests

The authors declare no competing interests.

Received: 30 December 2023 Accepted: 16 July 2024

Published online: 09 August 2024

#### References

1. Ishaq A, Asghar S, Gillani SA. Aspect-based sentiment analysis using a hybridized approach based on CNN and GA. *IEEE Access*. 2020;8:135499–512.
2. Kiritchenko S, Zhu X, Cherry C, et al. NRC-Canada-2014: Detecting aspects and sentiment in customer reviews. *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*. 2014: 437–442.
3. Jiang L, Yu M, Zhou M, et al. Target-dependent twitter sentiment classification. *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*. 2011: 151–160.
4. Sun Y, Wang Z, Zhang B, et al. Residents' sentiments towards electricity price policy: evidence from text mining in social media. *Resour Conserv Recycl*. 2020;160: 104903.
5. Miao Q, Li Q, Dai R. AMAZING: a sentiment mining and retrieval system. *Expert Syst Appl*. 2009;36(3):7192–8.
6. Chen P, Sun Z, Bing L, et al. Recurrent attention network on memory for aspect sentiment analysis. *Proceedings of the 2017 conference on empirical methods in natural language processing*. 2017: 452–461.
7. Trueman TE, Cambria E. A convolutional stacked bidirectional LSTM with a multiplicative attention mechanism for aspect category and sentiment detection. *Cogn Comput*. 2021;13:1423–32.
8. Zhou J, Huang JX, Chen Q, et al. Deep learning for aspect-level sentiment classification: Survey, vision, and challenges. *IEEE access*. 2019;7:78454–83.
9. Zhang S, Gong H, She L. An aspect sentiment classification model for graph attention networks incorporating syntactic, semantic, and knowledge. *Knowl-Based Syst*. 2023;275:110662.
10. Kim Y. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:14085882*. 2014. <https://doi.org/10.48550/arXiv.1510.03820>.
11. Zhou J, Jin S, Huang X. ADeCNN: an improved model for aspect-level sentiment analysis based on deformable CNN and attention. *IEEE Access*. 2020;8:132970–9.
12. Xiao L, Xue Y, Wang H, et al. Exploring fine-grained syntactic information for aspect-based sentiment classification with dual graph neural networks. *Neurocomputing*. 2022;471:48–59.
13. Gu X, Gu Y, Wu H. Cascaded convolutional neural networks for aspect-based opinion summary. *Neural Process Lett*. 2017;46:581–94.
14. Kuppusamy M, Selvaraj A. A novel hybrid deep learning model for aspect based sentiment analysis. *Concurr Comput*. 2023. <https://doi.org/10.1002/cpe.7538>.

15. Xue W, Li T. Aspect based sentiment analysis with gated convolutional networks. ArXiv preprint arXiv:180507043. 2018. <https://doi.org/10.48550/arXiv.1805.07043>.
16. Cuang F, Gao Q, Du J, et al., 2018, Convolution-based memory network for aspect-based sentiment analysis, The 41st International ACM SIGIR Conference pp. 1161–1164.
17. Phan HT, Nguyen NT, Hwang D. Convolutional attention neural network over graph structures for improving the performance of aspect-level sentiment analysis. *Inf Sci.* 2022;589:416–39.
18. Lin J, Najafabadi MK. Aspect level sentiment analysis with CNN Bi-LSTM and attention mechanism. *Inter J Sens Wire-less Commun Cont.* 2024;14(1):45–54.
19. Hinton GE, Krizhevsky A, Wang SD. Artificial neural networks and machine learning–ICANN2011. In: Honkela T, Duch W, Girolami M, Kaski S, editors. *Transforming auto-encoders*. Berlin: Springer, Berlin Heidelberg; 2011.
20. Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules. *Advances in neural information processing systems*, 2017, 30.
21. Qian Y, Wang J, Li D, et al. Interactive capsule network for implicit sentiment analysis. *Appl Intell.* 2023;53(3):3109–23.
22. Yang B, Han D, Zhou R, et al. Aspect opinion routing network with interactive attention for aspect-based sentiment classification. *Inf Sci.* 2022;616:52–65.
23. Zhao W, Ye J, Yang M, et al. Investigating capsule networks with dynamic routing for text classification. *ACL.* 2020. <https://doi.org/10.48550/arXiv.1804.00538>.
24. Geng R, Li B, Li Y, et al. Dynamic memory induction networks for few-shot text classification. arXiv preprint arXiv:200505727. 2020. <https://doi.org/10.48550/arXiv.2005.05727>.
25. Zhang X, Li P, Jia W, et al. Multi-labeled relation extraction with attentive capsule network. *Proc AAAI Conf Artific Intell.* 2019;33(01):7484–91.
26. She L, Gong H, Zhang S. An interactive multi-head self-attention capsule network model for aspect sentiment classification. *J Supercomput.* 2023;80(7):9327–52.
27. Wu Y, Guo X, et al. CharCaps: character-level text classification using capsule networks. *Intell Comput Technol Appl.* 2023;14087:187–98.
28. Wang J, Yu LC, Lai KR, et al. Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. *IEEE/ACM Trans Audio Speech Language Proc.* 2019;28:581–91.
29. Cambria E, Das D, Bandyopadhyay S, et al. *Affective computing and sentiment analysis*. In: Cambria E, Das D, Bandyopadhyay S, Feraco A, editors., et al., *A practical guide to sentiment analysis*. Cham: Springer International Publishing; 2017.
30. Yu J, Zha Z J, Wang M, et al. 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*. 2011: 1496–1505.
31. Ding X, Liu B, Yu PS. A holistic lexicon-based approach to opinion mining. *Proc 2008 Inter Conf Web Search Data Mining*. 2008. <https://doi.org/10.1145/1341531.1341561>.
32. Nguyen T H, Shirai K. Phrasernn: Phrase recursive neural network for aspect-based sentiment analysis. *Proceedings of the 2015 conference on empirical methods in natural language processing*. 2015: 2509–2514.
33. He R, Lee WS, Ng HT, et al. Exploiting document knowledge for aspect-level sentiment classification. arXiv preprint arXiv:180604346. 2018. <https://doi.org/10.48550/arXiv.1806.04346>.
34. Abdelgwad MM, Soliman THA, Taloba AI. Arabic aspect sentiment polarity classification using BERT. *J Big Data.* 2022;9(1):1–15.
35. Wang Y, Huang M, Zhu X, et al. Attention-based LSTM for aspect-level sentiment classification. *Proceedings of the 2016 conference on empirical methods in natural language processing*. 2016: 606–615.
36. Xu Q, Zhu L, et al. Aspect-based sentiment classification with multi-attention network. *Neurocomputing.* 2020;388:135–43.
37. Zhou J, Chen Q, Huang JX, et al. Position-aware hierarchical transfer model for aspect-level sentiment classification. *Inf Sci.* 2020;513:1–16.
38. Xu M, Zeng B, Yang H, et al. Combining dynamic local context focus and dependency cluster attention for aspect-level sentiment classification. *Neurocomputing.* 2022;478:49–69.
39. Xiao L, Xue Y, et al. Exploring fine-grained syntactic information for aspect-based sentiment classification with dual graph neural networks. *Neurocomputing.* 2022;471:48–59.
40. Zhang M, Zhang Y, Vo DT. Gated neural networks for targeted sentiment analysis. *Proc AAAI Conf Artific Intell.* 2016. <https://doi.org/10.1609/aaai.v30i1.10380>.
41. Parveen N, Chakrabarti P, et al. Twitter sentiment analysis using hybrid gated attention recurrent network. *Journal of Big Data.* 2023;10(1):50.
42. Kumar A, Vepa J. Gated mechanism for attention based multi modal sentiment analysis. *ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020: 4477–4481.
43. Liu N, Shen B. Aspect-based sentiment analysis with gated alternate neural network. *Knowl-Based Syst.* 2020;188: 105010.
44. Kumar A, Narapareddy VT, Srikanth VA, et al. Aspect-based sentiment classification using interactive gated convolutional network. *IEEE Access.* 2020;8:22445–53.
45. Ran X, Pan Y, Sun W, et al. Learn to Select via Hierarchical Gate Mechanism for Aspect-Based Sentiment Analysis. *IJCAI.* 2019: 5160–5167.
46. Lu Q, Zhu Z, Zhang G, et al. Aspect-gated graph convolutional networks for aspect-based sentiment analysis. *Appl Intell.* 2021;51(7):4408–19.
47. Han Y, Zhou X, Wang, et al. Fusing sentiment knowledge and inter-aspect dependency based on gated mechanism for aspect-level sentiment classification. *Neurocomputing.* 2023;551:126462.
48. Kamil G, Setiawan EB. Aspect-level sentiment analysis on social media using gated recurrent unit (GRU). *Build Inf Technol Sci (BITS)*. 2023;4(4):1837–44.
49. Ghorbanali A, Sohrabi MK. Exploiting bi-directional deep neural networks for multi-domain sentiment analysis using capsule network. *Multimed Tool Appl.* 2023;82(15):22943–60.

50. Su J, Yu S, Luo D. Enhancing aspect-based sentiment analysis with capsule network. *IEEE Access*. 2020;8:100551–61.
51. Xiaoxia Z, Xia Z. Attention based deep convolutional capsule network for hyperspectral image classification. *IEEE Access*. 2024. <https://doi.org/10.1109/ACCESS.2024.3390558>.
52. Wang Z, Shi-jie Hu, Liu W-D. Product feature sentiment analysis based on GRU-CAP considering Chinese sarcasm recognition. *Expert Syst Appl*. 2024;241: 122512.
53. Wang Y, Sun A, Huang M, et al. Aspect-level sentiment analysis using as-capsules. *World Wide Web Conf*. 2019. <https://doi.org/10.1145/33085583313750>.
54. Zhang B, Li X, Xu X, et al. Knowledge guided capsule attention network for aspect-based sentiment analysis. *IEEE/ACM Trans Audio Speech Language Proc*. 2020;28:2538–51.
55. Lin T, Sun A, Wang Y. EDU-capsule: aspect-based sentiment analysis at clause level. *Knowl Inf Syst*. 2023;65(2):517–41.
56. Tang D, Qin B, Feng X, et al. Effective LSTMs for target-dependent sentiment classification. *Comput Sci*. 2015. <https://doi.org/10.48550/arXiv:1512.01100v2>.
57. Gu S, Zhang L, Hou Y, et al. A position-aware bidirectional attention network for aspect-level sentiment analysis. *Proceedings of the 27th international conference on computational linguistics*. 2018: 774–784.
58. Song Y, Wang J, Jiang T, et al. Targeted sentiment classification with attentional encoder network. *Springer*. 2019;11730:93–103.
59. Tang H, Ji D, Li C, et al. 2020 Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification. *Proceedings of the 58th annual meeting of the association for computational linguistics*.: 6578–6588.
60. Li X, Bing L, Lam W, et al. 2018 transformation networks for target-oriented sentiment classification. *Proceedings of the 56th annual meeting of the association for computational linguistics*. 946–956.
61. Wu D, Wang Z, Zhao W. XLNet-CNN-GRU dual-channel aspect-level review text sentiment classification method. *Multimed Tools Appl*. 2023. <https://doi.org/10.1007/s11042-023-15026-4>.
62. Kannan G T, Gunasekar M, Ponnazhagan N A, et al. Aspect based sentiment aware word embedding for cross domain sentiment analysis. *2023 international conference on computer communication and informatics (ICCCI)*. *IEEE*, 2023: 1–5.
63. Chen Z, Qian T. Transfer capsule network for aspect level sentiment classification. *Proceedings of the 57th annual meeting of the association for computational linguistics*. 2019: 547-556.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.