DOI: 10.54254/2755-2721/6/20230417

CNNs and RNNs in aspect-level sentiment analysis and comparison

Qi Yijun

Engineering, The University of Leicester, Leicester, LE1 7RH, United Kingdom

yq63@student.le.ac.uk

Abstract. Aspect-based sentiment analysis (ABSA), also known as fine-grained sentiment analysis, may offer a precise polarity for each aspect in statement aspect. CNN and RNN neural network models, the most fundamental and widely used deep neural network models, have been created by researchers using a variety of methodologies to offer reliable findings in ABSA. This paper will sort out the development process of CNN and RNN models in completing various tasks in aspect-level sentiment analysis and find out the current SOTA model. In addition, by comparing the representatives of the two models in the same environment, the advantages and disadvantages of the two models are analysed to provide guidance for choosing different baseline models. The results show that the CNN-based model has the advantages of high accuracy and fast training speed, but the design is more complicated, and the pooling layer is prone to lose position or sequence features. The RNN-LSTM based model is easier to capture sequence data, has high reliability, and has a relatively simple design structure. However, it requires significantly longer training time, the model is complex, and the computational cost is high. Considering the characteristics of the two most basic neural models, new tasks require researchers to propose better hybrid neural network models.

Keywords: deep learning network models, CNN, RNN, sentimentanalysis, aspect-level sentiment.

1. Introduction

With the increasing use of mobile Internet nowadays, it has become customary for people to use network platforms to disseminate information, provide viewpoints, and express feelings. On social media sites, users post a lot of emotionally charged texts that reflect their personal opinions and levels of interest in the topics at hand. The rapid expansion of numerous e-commerce platforms has also increased the size, digitality, and openness of text information [1]. As a result, big platforms started to gather, arrange, and analyze textual data that contained emotional inclinations in order to comprehend social hotspots, product trends, user activity data, etc. It is also significant for early warning and managing public opinion on the Internet [2]. It has been a big subject in the academic community to figure out how to efficiently and reliably analyze the vast amounts of unstructured data present in social networks and assess users' emotional inclinations. As a result, sentiment analysis technology is quite important for study.

Determine the sentiment category of texts with sentimental inclinations is the core objective of the Sentiment Analysis (SA) task, which is a part of natural language processing. It may be categorized into three distinct of sentiment analysis tasks, including document-level, sentence-level, and aspect-level tasks, depending on the amount of granularity [3]. Aspect Based Sentiment Analysis (ABSA) is a more in-depth sentiment analysis based on the target in the corpus, in contrast to the preceding two coarse-grained research techniques. The aspect word may consist of one or more words. Three phases are involved in aspect-level sentiment analysis: aspect word recognition, extraction, and sentiment categorization. The primary one is sentiment classification [4].

The study of aspect-level sentiment analysis using deep neural network models has recently gained popularity in academia due to the unexpected rise of deep learning. The long short-term memory model (LSTM), which is typically regarded as one of the RNN models, and the convolutional neural network (CNN) are two of the numerous deep learning network models that are now more mature, frequently used, and better appropriate for this job [5]. Compared to sentiment analysis methods based on sentiment dictionaries and traditional machine learning, the neural network method has significant advantages in text feature learning because it can actively learn features and retain the information of the words in the text, allowing for better sentiment analysis [6].

The author of this study evaluates the development of CNN and RNN models on the problem of Aspect-based sentiment analysis, respectively. We will see their infrastructure, how they have evolved, and what progress has been made recently. The author will then examine some recent success stories and compare them to demonstrate their respective strengths and weaknesses.

Then in Section 2, the author survey CNN-based models, including their models, development process, and some recent solutions. In Section 3, RNN-based models are studied from the same perspective. In section 4, several state-of-the-art representative models are selected for comparative analysis in the same environment, with the author's own views. Section 5 summarizes the views of this paper.

2. CNN-based methods

2.1. Introduction of CNN

One of the critical deep learning network models is CNN. CNN has made considerable strides in several domains since Hubel and Wiesel's [7] 1960s proposal. An increase in the text sentiment analysis field using CNN has been seen in recent years, with powerful results on several standard datasets. The greatest semantic feature in the text is often used to classify text sentiment using CNN. When initialization, the convolution layer is used to extract the sentence's features, and the maximum pooling layer is used to choose the features' most relevant jobs. CNN's primary benefit is its capacity to extract the most significant n-gram features and turn them into an "informative latent semantic representation" for further categorization [1]. Figure 1 shows a four-layer base CNN model that is taken from a study by Hai HaDo [1].

2.2. Architecture

According to Kim [4], the basic CNN for sentiment analysis may include four layers, as shown in Figure.1. In this context, the convolution layer and the max pooling layer of CNN have two important implications: the convolutional layer can keep the most essential part of each feature map from being lost, and the highest pooling layer can make the output length fixed and not affected by the filter window size [3]. The adoption of a CNN model in the ABSA study can extract local patterns inside data no matter where they are because their design allows them to find these features [8]. Figure.2 is an example of the CNN method from the paper of Hai HaDo et al. [1].

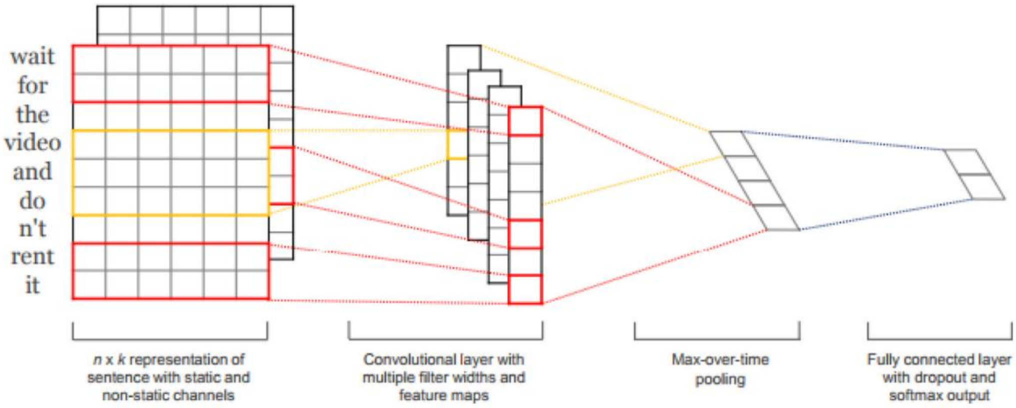

Figure 1. 4 layers CNN models.

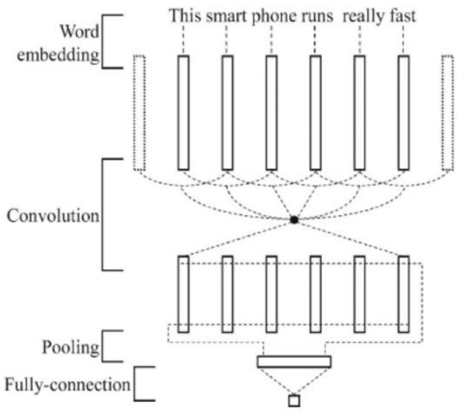

Figure 2. An example of CNN architecture.

2.3. *Evolutionary process and SOTA model*

2.3.1. Evolutionary process. An example of effective research for the ABSA subtask "Opinion target extraction" is the adaptation of a CNN model from sentence encoding to phrase prediction [9]. They create a local characteristic window of five words surrounding each word in the phrase, assuming that the label of each word relies on each neighboring word. Each word window is then subjected to a seven-layer deep CNN. Experiments demonstrate that deep CNN models still beat some SOTA models even in the absence of feature engineering or linguistic patterns.

Additionally, CNN is a promising method for various ABSA jobs. Toh and Su assembled two separate machine learning algorithms in paper to attain the most significant result in ACD's SemEval 2016 [4]. They utilized a binary correlation strategy since they approached ACD as a multi-class classification issue. To determine whether text comprised an aspect category, they specifically integrated the probabilities produced by numerous classification algorithms trained on a single-layer feedforward neural network. CNN characteristics have the most impact on performance as compared to other features.

Other studies using multi-task CNNs demonstrate the extra benefits that CNNs may offer. In Gu's [10] study, examples of several CNNs are shown. They propose a cascaded system consisting of a sentiment classifier and an aspect mapper, both using CNN models. Two CNNs are arranged in a cascade. Each mapper evaluates the input sentence to see if it corresponds to that aspect. If so, the sentiment classifier forecasts whether the emotion is good or negative. The cascaded model demonstrates that

CNN greatly decreases elapsed time when compared to SVM, in addition to the advantage of decreasing feature engineering when compared to conventional ML approaches.

However, CNN models are still not very good at modeling sequence information and capturing long-range relationships. Since then, many academics have started to work on overcoming this constraint, such as Xue and Li [11], who suggested a model based on convolutional neural networks and the gating mechanism and aspect embedding (GCAE) instead of using attention mechanisms. The model is more precise and productive. Ren et al. presented the distillation network (DNet), a compact and efficient emotion evaluation model based on gated convolutional neural networks [4]. The model first encrypts the information in a phrase to identify which information helps predict sentiment polarity. Then they retrieve aspect-sensitive data by using the GCAE's gating unit. In Table 2, the performance of several ABSA models based on CNN are given.

Table 1. The performance of some CNN-based ABSA methods.

	Accuracy (%)		Results(F1) (%)	
Term	Laptop	Restaurant	Laptop	Restaurant
PosATT-GTRU(2021)	83.38	76.49	\	\
RPAEN(2020)	74.1	81.2	70.5	72.5
RPAEN-BERT(2020)	80.6	85.2	76.8	78.0
Dnet-GTU(2020)	73.98	79.38	68.28	68.12
GCAE(2018)	69.4	77.28	\	\

2.3.2. An example of a SOTA model. Ramaswamy S. L.'s RecogNet-LSTM+CNN model outperforms other models in classifying aspects and opinions [12]. The researchers substantially improved the efficiency of aspect extraction and classification efficiency by combining explicit sentiment and semantic signals from the external RecogNet knowledge base with implicit data from the LSTM model. To further increase classification accuracy, a CNN with object and location attention mechanism is also built on the RecogNet-LSTM layer [12]. This RecogNet-LSTM+CNN model with attention mechanism outperforms other CNN-based hybrid models in terms of aspect classification and opinion classification.

3. RNN-based methods

3.1. Introduction

CNNs have the apparent drawback of being unable to handle continuous data input, even though they can provide effective results. RNNs, which are neural networks with internal loops that can store data, are ideally suited to tackling the issue of continuous data input in order to resolve the emotion categorization problem. By using recursive computing, the model learns feature vectors for lengthy texts. However, RNNs have a gradient vanishing problem when dealing with long-distance dependencies, which makes RNNs unable to learn long-distance dependency information. To solve this problem, the Long Short-Term Memory Model (LSTM) was derived [5]. LSTM has better memory capacity and can make effective use of contextual feature information as well as the ability to match nonlinear relationships, maintaining text ordering information. The LSTM-based model derived from RNN is the most often utilized model in the ABSA job.

There is also the Gated Recurrent Unit (GRU). GRUs are a type of LSTM with more excellent durable memory, making them ideal for capturing long-term relationships between sequence parts [3]. Because GRU models are simpler than LSTMs, numerous researchers have begun to use them to build ABSA problems.

3.2. Architecture

3.2.1. LSTM. Compared with the standard feedforward neural network, LSTM has the characteristics of feedback link. This allows it to handle more complex complete sequences of data [5]. The LSTM architecture has three gates, the input gate adds data to the cell state, the forget gate removes data from the cell state, and the outputs select important information from the current state and display it. A typical Architecture of standard LSTM model is shown in Figure 3 [13].

3.2.2. GRU. The GRU model, based on LSTM, merges the input gate and forget gate into an update gate. Along with other adjustments, it also takes cell state and concealed state into account. The final model is less complex than typical LSTM models. Reset and update gates exist in GRUs. While the latter is in charge of deciding how much of the previous hidden state should be carried over to the next state, the former has the option to diminish the hidden state altogether if it thinks that the prior hidden state is unimportant to the computation of the current state [3]. A typical Architecture of GRU model is shown in Figure 4 [3].

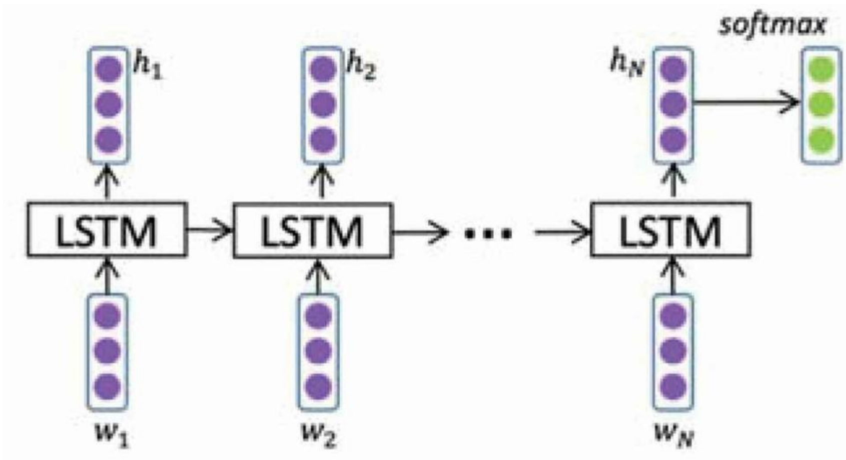

Figure 3. Architecture of standard LSTM.

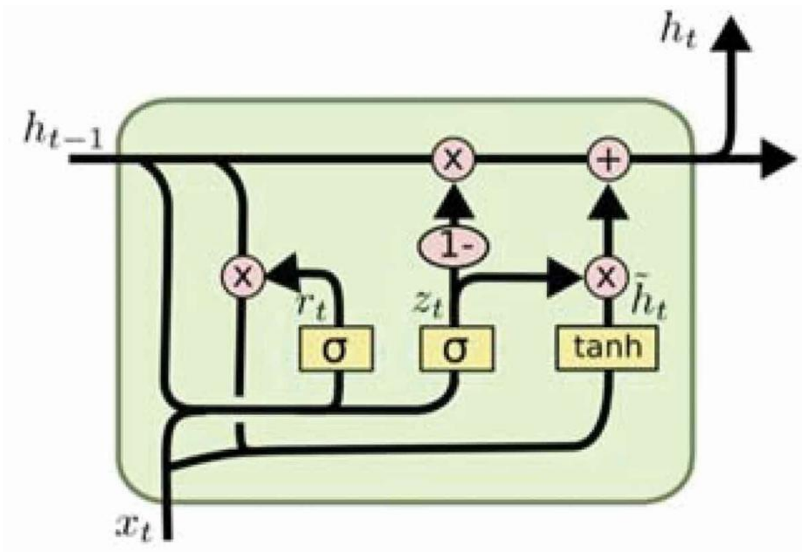

Figure 4. Gated Recurrent Unit Neural network.

3.3. *Evolutionary process and SOTA model*

3.3.1. Evolutionary process. Most researchers employ LSTM techniques for sentiment analysis tasks, particularly in ABSA, since they can capture sequence models. Target-dependent sentiment classification (TD-LSTM) was proposed in paper [14]. In order to solve the problem that the ordinary LSTM model cannot collect target information, this model uses two LSTM models, one left and one right, to model the context and place the target in the middle to make predictions for multiple targets in a phrase.

Some researchers proposed a new idea of adding aspect word embeddings to LSTMs (ATAE-LSTM) [13]. In addition to adding aspect embeddings, they employ an attention mechanism as well. The ABSA assignment will now use aspect terms for the first time. Some researchers have proposed the Interactive Attention Network, a model based on LSTM method combined with attention mechanism, which can learn attention in context and target while generating representations for target and context [15]. With the ATAE-LSTM model, the research started modeling different objects differently. Researchers began to model aspects and texts simultaneously, a departure from earlier research ideas.

The joint attention LSTM network (JAT-LSTM) proposes to create a joint attention LSTM network by fusing aspect attention with emotion attention [16]. The model enhances the richness of the input information to the LSTM network by combining aspect word embedding, sentiment word embedding, and sentence embedding.

The researchers then designed a target-related sentiment analysis model (TD-BiGRU) by adopting the bi-GRU model to extract targets and assess sentiment polarity. The model can find and extract targets from the Twitter database, reflect the relationship between the target and its context, and determine the polarity of the tweet where the target is located [17]. Subsequently, the researchers proposed a GRU-based bidirectional position-aware attention network (PBAN) [18]. It not only considered the position information of aspect-level words but also adopted a bidirectional attention mechanism to represent the relationship between words and sentences. Relationship. However, most of these models ignore the location information of aspects when encoding phrases. To address this issue, HAPN was proposed [19]. When modeling phrases, they add positional embeddings and go on to produce position-aware representations. The findings demonstrate that the technique performs significantly better than the SOTA methods for aspect-level sentiment categorization. In Table 2, the performance of several ABSA models based on LSTM and GRU are given.

Table 2. The performance of some LSTM-based and GRU-based ABSA methods.

	Accuracy(%)	
Term	Laptop	Restaurant
BEBA (2022)	84.55	77.42
JAT-LSTM (2018)	80.5	88.3
ATAE-LSTM (2016)	68.7	77.2
ABAE-Bi-GRU (2019)	71.2	81.2
HAPN (2018)	77.27	82.23

3.3.2. An example of SOTA models. LIN M et al. proposed an aspect sentiment that combines BERT embedding and attention mechanism analytical model to address the issues of low text utilization, difficulty extracting helpful information, and inability to efficiently identify word ambiguity in aspect sentiment analysis task research [2]. The model builds a BERT-BiLSTM-Attention architecture for describing phrase sentiment by fusing the BERT pre-trained language model with the multi-head attention mechanism. Firstly, the BERT model with a bidirectional Transformer structure is used to generate semantic vectors according to the word context dynamically; then, the semantic representation of text words is enhanced by effectively identifying the ambiguity of words, and the semantic vectors are input into BiLSTM to capture the bidirectional semantic dependencies. The model's ability to extract practical information; finally, combined with the multi-head attention mechanism with parallel information processing ability to parse the encoded text representation and obtain information with a high degree of emotional relevance [2]. The results in the dataset show that the model has the best accuracy compared to the baseline model. Figure 5 shows a standard BEBA model structure [2].

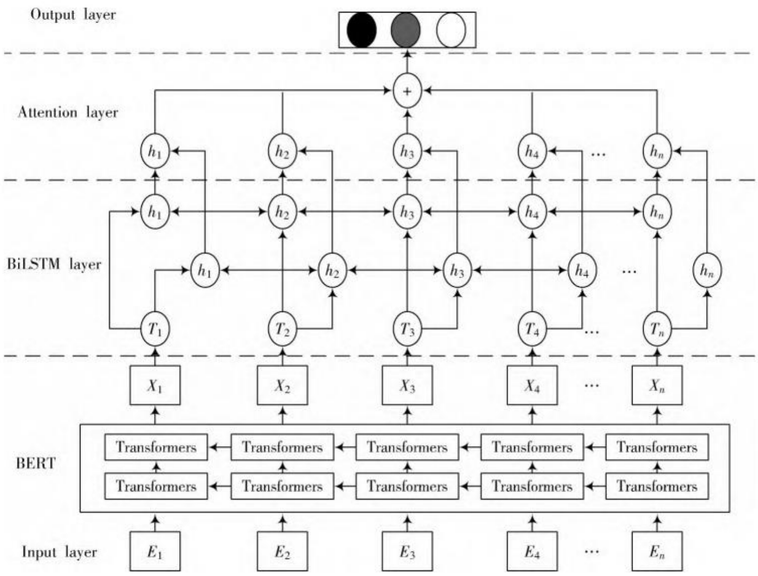

Figure 5. BEBA model structure diagram.

4. Comparison and analysis

In order to compare the ABSA models based on the two neural network models, the author uses the comparative experiment method to compare the models in the literature under the same experimental environment and experimental data. The public data set SemEval 2014 is selected for simulation experiments. There are two types of reviews in the dataset. One is a dataset in the laptop domain, and the other is a dataset related to the restaurant domain. Each comment information can have multiple aspect words, and each has three polarity options: positive, negative, and neutral. In the experiments in this paper, the data is preprocessed concerning typical methods, the "conflict" category in the sentiment polarity category of the original dataset is removed, and only "positive," "negative," and "neutral" are retained. Finally, 4728 and 2966 review data from restaurants and Laptops were used. Table 3 and Table 4 show the statistical results of sentiment categories in Restaurant and Laptop datasets, respectively.

Table 3. Category statistics table for the restaurant dataset.

Dataset	Positive	Negative	Neural	Total
Restaurant(train)	2064	707	537	3508
Restaurant(test)	828	296	296	1220

Table 4. Category statistics table for the laptop dataset.

Dataset	Positive	Negative	Neural	Total
Laptop(train)	894	770	564	2228
Laptop(test)	441	228	269	538

Experiment was carried out with different model methods on the same dataset, running ten times to take the average as the final result. In addition, the same server was used to run each comparison model to calculate the running time, and the average running time was obtained after running 30 epochs for comparison. Table 5 shows the accuracy results of the three models. Table 6 shows the runtimes of the three models in the dataset.

DOI: 10.54254/2755-2721/6/20230417

Table 5. Accuracy results of each model / (%).

Terms	Restaurant	Laptop
PosATT-GTRU	81.38	76.49
BEBA	80.55	77.42
HAPN	77.27	82.23

Table 6. Comparison of model runtimes in the dataset / (s).

Terms	Restaurant	Laptop
PosATT-GTRU	5.11	2.52
BEBA	125.41	69.86
HAPN	4.23	2.36

It can be seen that the use of LSTM network in the model will increase the running time of the model, and if the model is built with multiple layers or even more complex LSTMs, the time complexity of the model will be seriously increased. In addition, when the convolution operation is used, the model consists of multiple computing layers, and the operation of text information from shallow to deep layers takes more time, so this model consumes a relatively large amount of time in models other than the LSTM model. In general, using the gating mechanism in the model can effectively control the flow of input information, optimize the running time of the model, and improve the efficiency of sentiment classification.

The data shows that the CNN-based models have the advantages of high accuracy and fast training. However, they are more complicated to design, and the pooling layer is prone to losing positional or sequential features. The RNN (LSTM)-based model is easier to capture sequential data, has high reliability, and has a relatively simple design structure. However, it requires significantly longer training time, the model is complex, and the computational cost is high.

Research in the field of ABSA is gradually developing towards multitasking learning, and efforts are being made to explore effective interaction mechanisms between tasks. Taking the characteristics of the two most basic neural models into account, the new task requires researchers to propose a better hybrid neural network model. There are already examples of this in sentiment analysis. However, the exploration in fine-grained ABSA is still sparse.

5. Conclusion

In this article, the author investigate ABSA tasks. First, the author describe the task of ABSA and a solution based on two standard deep learning network models. The solution based on the CNN model has the advantages of high accuracy and fast training speed. However, the design is more complicated, and the pooling layer is prone to losing positions. The solution based on RNN, especially LSTM, is the most common. Its model is more straightforward and reliable, but the training time is long, and the design is more complicated. By comparing several SOTA models in the same environment, the author summarizes the advantages and disadvantages of the two models and provides a basis for choosing different models to solve tasks with different needs in different situations. However, research in the field of ABSA is gradually developing towards multitasking, and researchers are working hard to explore the interaction between tasks. Building a high-accuracy university model in the ABSA task is still very challenging.

In addition, the application of more basic deep learning network models in tasks is still to be developed, such as graph neural network models. Moreover, newer models can be considered based on these models in the future. It can also be seen from the survey that each dataset has different methods to achieve better performance. And future research should focus on finding solutions that can be flexibly used in multiple datasets.

References

[1] Do H H, Prasad P W C, Maag A and Alsadoon A 2019 Deep learning for aspect-based sentiment analysis: a comparative review. J. Expert systems with applications, 118, 272-299.
[2] Lin M, Li S 2022 Aspect sentiment analysis integrating BERT embedding and attention mechanism. J. Modern Electronics Technique. 45(12):130-136.
[3] Komang W T and Huang J J 2022 Deep Learning Approach for Aspect-Based Sentiment Classification: A Comparative Review. J. Applied Artificial Intelligence, 36:1.
[4] Ren F, Feng L, Xiao D, Cai M and Cheng S 2020 DNet: a lightweight and efficient model for aspect based sentiment analysis J. Expert Systems with Applications 151, 113393.
[5] Sherstinsky A 2020 Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. J. Physica D: Nonlinear Phenomena, 404, 132306.
[6] Toh Z and Su J 2016 Nlangp at semeval-2016 task 5: Improving aspect based sentiment analysis using neural network features. In Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016) (pp. 282-288).
[7] Hubel D H and Wiesel T N 1962 Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. The Journal of physiology, 160(1), 106.
[8] Beini, Z, Xuee, C, Bo, L, and Weijia, W 2021. A new few-shot learning method of digital PCR image detection. IEEE Access, 9, 74446-74453.
[9] Dashtipour K, Poria S, Hussain A, Cambria E, Hawalah A Y, Gelbukh A and Zhou Q 2016. Multilingual sentiment analysis: state of the art and independent comparison of techniques J. Cognitive computation, 8(4), 757-771.
[10] Gu X, Gu Y and Wu H 2017 Cascaded convolutional neural networks for aspect-based opinion summary. J. Neural Processing Letters, 46(2), 581-594.
[11] Xue W and Li T 2018 Aspect based sentiment analysis with gated convolutional networks arXiv preprint arXiv:1805.07043.
[12] Ramaswamy S L and Chinnappan J 2022 RecogNet-LSTM+ CNN: a hybrid network with attention mechanism for aspect categorization and sentiment classification. J. Journal of Intelligent Information Systems, 58(2), 379-404.
[13] Wang Y, Huang M, Zhu X and Zhao L 2016 Attention-based LSTM for aspect-level sentiment classification. In Proceedings of the 2016 conference on empirical methods in natural language processing (pp. 606-615).
[14] Tang D, Qin B, Feng X and Liu T 2015 Target-dependent sentiment classification with long short term memory arXiv preprint arXiv:1512.01100.
[15] Ma D, Li S, Zhang X and Wang H 2017 Interactive attention networks for aspect-level sentiment classification. arXiv preprint arXiv:1709.00893.
[16] Cai G and Li H 2018 Joint attention LSTM network for aspect-level sentiment analysis. In China Conference on Information Retrieval (pp. 147-157).
[17] abreel M, Hassan F and Moreno A 2018 Target-dependent sentiment analysis of tweets using bidirectional gated recurrent neural networks. In Advances in hybridization of intelligent methods (pp. 39-55).
[18] Gu S, Zhang L, Hou Y and Song, Y 2018 A position-aware bidirectional attention network for aspect-level sentiment analysis. In Proceedings of the 27th international conference on computational linguistics (pp. 774-784).
[19] Li L, Liu Y and Zhou A 2018 Hierarchical attention based position-aware network for aspect-level sentiment analysis. In Proceedings of the 22nd conference on computational natural language learning (pp. 181-189).