Computer Science and Information Systems 20(1):77–93 https://doi.org/10.2298/CSIS220303047X

SEE-3D: Sentiment-driven Emotion-Cause Pair Extraction Based on 3D-CNN*

Xin Xu[1,2], Guangli Zhu[1,2†], Houyue Wu[1,2], Shunxiang Zhang[1,2], and Kuan-Ching Li[3]

[1] School of Computer Science and Engineering,
Anhui University of Science & Technology, 232001 Huainan, China
[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center,
WangJiang Road 5089, Hefei, 230088, Anhui, China
xuxin.onepiece@gmail.com
glzhu@aust.edu.cn
why3664@163.com
sxzhang@aust.edu.cn
[3] Department of Computer Science and Information Engineering (CSIE),
Providence University, 43301 TaichungTaichung, Taiwan
kuancli@pu.edu.tw

Abstract. As an emotional cause detection task, Emotion-Cause Pair Extraction (ECPE) provides technical support for intelligent psychological counseling, empty-nest elderly care, and other fields. Current approaches mainly focus on extracting by recognizing causal relationships between clauses. Different from these existing methods, this paper further considers the influence of sentimental intensity to improve extraction accuracy. To address this issue, we propose an extraction model based on sentiment analysis and 3D Convolutional Neural Networks (3D-CNN), named SEE-3D. First, to prepare fundamental data for sentiment analysis, emotion clauses are clustered into six emotion domains according to six emotion types in the ECPE dataset. Then, a pre-trained sentiment analysis model is introduced to compute emotional similarity, which provides a reference for identifying emotion clauses. In the extraction process, similar features of adjacent documents in the same batch of samples are fused as input of 3D-CNN. The 3D-CNN enhances the macro semantic understanding ability of the model, thereby improving the extraction performance. The results of experiments show that the accuracy of ECPE can be effectively improved by the SEE-3D model.

Keywords: ECPE, Sentiment analysis, Neural networks, 3D-CNN.

---

* The initial idea has been presented with some preliminary experimental results in ATCI 2022 [1]. The current paper is an extended version that contains some new ideas, formulations, and more extensive experimental results.

† Corresponding author

1. Introduction

With the accumulation of life pressure and the accelerated pace of life, people need more emotional understanding than ever before. Different from traditional psychological counseling services, ECPE 2 can provide technical support for intelligent psychological counseling, empty-nest elderly care, and other fields. However, although many excellent methods have emerged in recent years, the extraction effect can still be improved if the sentimental intensity and template similarity are considered.

In the related research on sentiment analysis, we propose a key sentence extraction algorithm and a sentiment classification model (SC-CMC-KS) for processing microblog comment texts 3. Combined with a hybrid neural network and ELECTRA, a sentiment classification model 4 is proposed for Chinese short comment texts. In 5, Xu et al. proposed an emotional element extraction model (ALSEE) for extracting keywords from all aspects of product reviews. Unlike our previous works, sentiment analysis can also be utilized for relation extraction tasks at the clause level. To further apply sentiment analysis to the ECPE task, the following two aspects need to be considered: (1) The dataset contains many clauses which need to be divided according to the type of emotion. The intensity of emotion needs to be quantified by a unified indicator. (2) To mine the similarity of documents, it is necessary to build candidate emotion-cause pairs (EC-Pairs) and construct a network to realize parallel processing of multiple documents.

Based on the above two considerations, we constructed an ECPE model based on sentiment analysis and 3D-CNN, named SEE-3D. The process of extracting EC-Pairs can be divided into three parts: encoding clauses as features, sentiment analysis & classification, and 3D fusion extraction. The framework of SEE-3D is illustrated in Figure 1:

(1) **Encoding Clauses as Features**. The word embedding is generated by BERT, which can carry rich semantic information. A word-level BiLSTM is constructed to encode semantic information of context words. To further obtain the intrinsic relationship between clauses, we introduced the attention mechanism to BiLSTM and encoded clauses as features at the coarse-grained level.

(2) **Sentiment Analysis & Classification**. The features generated in the previous step are the original input of this part. Firstly, each clause in the same document is paired with each other as EC-Pairs, and a quantitative method based on Euclidean distance is proposed as an emotional similarity. Then, the classification module concatenates the clause vectors of EC-Pairs, a linear layer is introduced to calculate the probability. By calculating the emotional similarity of emotion clauses in candidate EC-Pairs, we further obtain the emotional similarity of all EC-Pairs. Finally, to fuse emotional information and probability, each document is constructed as a probability matrix in the form of a 2D matrix.

(3) **3D Fusion Extraction**. After analyzing the documents in the dataset, we hypothesize template similarity between documents. To further prove this hypothesis, we stack 2D matrices constructed in the previous part into 3D matrices and pass them into a 3D-CNN for fusion extraction. After performing a 3D convolution operation, the network's output is split into several 2D probability matrices, i.e., the final extraction result is obtained. Ablation experiments in section 4.5 proved the validity of the hypothesis.

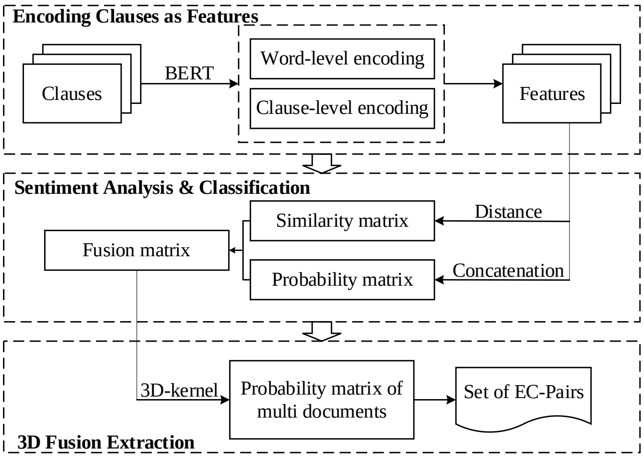

Fig. 1. System Framework of SEE-3D

In this paper, to further represent the sentimental intensity of clauses, an emotional similarity representation method is proposed for the ECPE task. We apply Euclidean distance to measure the space between clauses and the center of the emotion domains. In brief, the closer the distance to the emotion domain is, the greater the probability of the emotion clause is. We incorporate sentimental intensity into the EC-pairs extraction process through this similarity representation. To enhance the understanding ability of the model, two prediction probabilities of documents are fused by a 3D fusion extraction method. Thus, the SEE-3D model can reach relatively high accuracy and significantly improve emotion clause extraction accuracy.

The SEE-3D model is trained by using the ECPE benchmark dataset. The advantage of the encoding module is that the semantic information carried by BERT can be learned. The sentiment analysis module is introduced as a pre-trained model, providing a reference for representing emotional similarity. Meanwhile, 3D-CNN provides strong support for capturing local semantic information. The advantage of the SEE-3D model is that it considers the sentimental intensity of clauses and improves the accuracy of EC-Pairs extraction.

In this paper, the remaining sections are organized as follows. In Section 2, we introduce previous research on ECPE; Section 3 introduces the SEE-3D model by module; Analyzes and discusses the results of experiments are shown in Section 4; Finally, future work and summary are presented in Section 5.

2. Related Work

ECPE is developed from ECE 6. The task goal of ECE emphasizes finding the cause of the emergence of an emotional word, i.e., locating the corresponding cause clause through emotional words. In 7, Xia et al. proposed a hierarchical RNN method for ECE. The task goal of ECPE emphasizes finding out which events are the causes of an emotional event, i.e., matching the corresponding cause clauses through emotion clause positioning.

2.1. From ECE to ECPE

ECE is based on word-level, focusing on detecting cause events. In 8, Ding et al. combined the three features of relative position, text content, and global label to improve the accuracy of ECE extraction. In recent years, although there have been many excellent research results 9101112, limited by the utilization of emotional words, the research results are not easy to apply to realistic scenarios. ECPE is based on clause-level, focusing on causal event pairing. In 2, Xia et al. split the extraction task into two sub-steps, the first classifying the clauses into sentiment or reason clauses and the second calculating all possible EC-Pairs and filtering them with custom filtering rules.

2.2. Recent research on ECPE

To avoid the error propagation caused by two-step extraction, In 13, Song et al. proposed an end-to-end network based on connection learning, which represents clauses as nodes and extracts EC-Pair through directed connectivity learning between nodes, while the study 14 proposed a range control mechanism to discover the interaction between cause and emotion through multi-task learning. In 15, Tang et al. used a multi-layer attention mechanism to model the intrinsic semantic information in EC-Pairs. Study 16 proposed a classification method that did not depend on the extraction results and improved the model's performance by adding a position awareness mechanism. In 17, Chen et al. constructed a pair graph to represent the relationship pairs in the document,three kinds of dependencies were defined, and the graph convolution neural network was utilized to learn the dependencies for EC-Pairs extraction. Study 18 regarded ECPE as a sequence labeling task and used CNN and BiLSTM to extract emotional cause pairs after the unified labeling of data sets. In 19, Cheng et al. designed an asymmetric network to find the local area EC-Pair by cross-subnetwork local pair searcher. Study 20 proposed a joint framework for multi-task learning and used the sliding window mechanism to reduce the noise generated by long-distance clauses during extraction. In 2122, Yuan et al. and Fan et al. respectively proposed two labeling strategies based on different characteristics of the ECPE task to improve accuracy. By analyzing the construction process of a directed graph with labels, Fan et al. captured the causal relationship between clauses 23. In addition, some end-to-end models 242526 and two-step models 2728 have achieved good performance.

Although ECPE does not need to annotate emotional words in advance, the influence of emotional words on the extraction results cannot be ignored. Therefore, we propose the SEE-3D model in this paper and integrate emotional distance as an essential indicator in the extraction process.

3. Methods

3.1. Architecture

The SEE-3D model consists of three modules: (a) Encoding Clauses as Features; (b) Sentiment Analysis & Classification; (c) 3D Fusion Extraction. The overall structure is shown in Figure 2.

In Figure 2(a), the model regards each clause in documents as a word sequence $c = \{w_1, w_2, ..., w_{|c|}\}$, which w represents a word in the clause and $|c|$ is the number of words, and generates clause-level representation by BiLSTM with an attention mechanism. Figure 2(b) contains two sub-modules:

1. The pre-trained sentiment analysis model aims to convert the output of Encoder into a sentiment vector. By calculating the Euclidean distance between the sentiment vector and the center of the nearest emotion domain, we convert the distance into similarity to judge the probability of an emotion clause. Euclidean distance can be more accurate in calculating the same dimension vector in K-means clustering space. The parameters of the sentiment analysis model are frozen during training.
2. The classification module regards all clauses as emotion and cause to build emotion set and cause set, then construct candidate EC-Pairs as $ECPs = \{(e^1, c^1), (e^2, c^2), ..., (e^m, c^m)\}$, where m is the maximum number of clauses for all documents. For each (e, c) in ECPs, pass $r^e \oplus r^c$ into a linear layer to calculate the probability that (e, c) is true, where r^e and r^c is the output of e and c from BiLSTM in Figure 1.

The input of Figure 2(c) is a three-dimensional matrix, and each element of the matrix is the output weighted sum of the two sub-models in Figure 2(b). 3D-CNN learns the global similarity between samples, and the extraction results of each document are finally output. The red lattice in Figure 2(c) represents the EC-Pair with the highest prediction probability of the model as the final extraction result.

3.2. Emotion Domain

The ECPE benchmark corpus 29 contains six emotions. Our analysis of samples in the corpus can be concluded as follows: the polarity and strength of the six emotions are significantly different, and they are pretty different from the clauses without emotional words. The distribution of emotion types of samples is shown in Table 1. Sadness accounted for the highest proportion of 26.94% of the six emotions, and Surprise accounted for the lowest proportion of 4.18%.

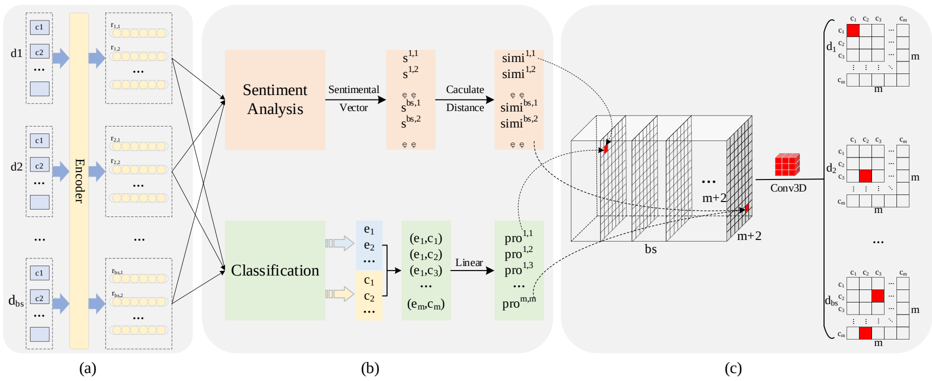

Fig. 2. Architecture of SEE-3D

Table 1. Distribution of emotion types

Emotion	Percentage	Number
Surprise	4.18%	88
Disgust	10.69%	225
Anger	14.35%	302
Fear	18.00%	379
Happiness	25.83%	544
Sadness	26.94%	567

We hope to provide a reference for verifying whether clauses contain emotions by constructing six types of emotion domains. Firstly, a pre-trained sentiment analysis model 4 is utilized to convert each clause-level vector into a sentiment vector. Then all sentiment clauses in the corpus are clustered by a k-means 30 clustering algorithm (k=6) to construct the emotion domains of six sentiment types, denote as $domain^S = \{sd_1, sd_2, ..., sd_6\}$. In the follow-up work, we obtain the cluster centers of six emotion domains, denote as $domain^C = \{sdc_1, sdc_2, ..., sdc_6\}$, where each represents the center of emotion domains. By calculating the center vector and radius of each cluster, it can provide support for subsequent emotional computing. Based on the above operations, the emotional distance between clauses and emotion domains can be obtained directly.

3.3. Encoding Clauses as Features

The input of the model is one batch of documents is denoted as $D = \{d_1, d_2, ..., d_{bs}\}$, where bs is the batch size of documents. For each document d_i in D, it contains several clauses $\{c^{i,1}, c^{i,2}, ..., c^{i,|d_i|}\}$, where |di| is the number of clauses in d_i. Any clause $c^{i,j}$ is regarded as a word sequence: $\{w_1^{i,j}, w_2^{i,j}, ..., w_{|c^{i,j}|}^{i,j}\}$. Each clause is passed into BiLSTM to encode semantic information between the words and obtain the hidden state of BiLSTM $h^{i,j} = \{h_1^{i,j}, h_2^{i,j}, ..., h_{|c^{i,j}|}^{i,j}\}$.

$$\alpha_k = \frac{\exp(H_k^{i,j} W_k)}{\sum_{|c_{i,j}|} \exp(H_k^{i,j} W_k)} \tag{1}$$

$$r_j^i = \sum_{k=1}^{|c_{i,j}|} \alpha_k H_k^{i,j} \tag{2}$$

where W_k is a learnable weight parameter. Figure 3 describes the process of generating clause-level representations from a document. The details of BiLSTM 31 and attention mechanism 32 are omitted in this paper.

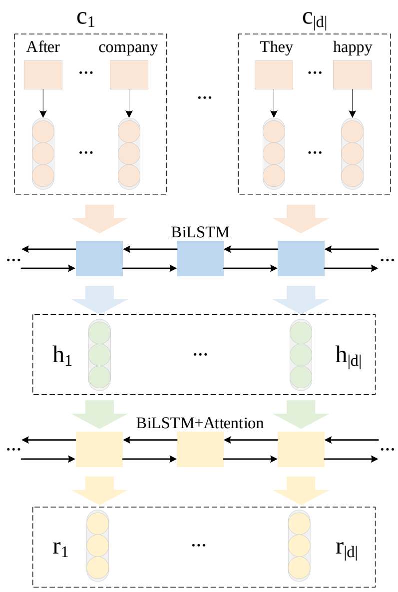

Fig. 3. Process of encoding a document

3.4. Encoding Clauses as Features

This part consists of two subtasks: (1) Generate the sentiment vectors of the clauses and calculate the distances from emotion domains; (2) Regard each clause as both cause and emotion, then construct all candidate EC-Pairs in a document.

The pre-trained sentiment analysis model generates the sentiment vector $s^{i,j}$ for the j-th clause $c^{i,j}$ of the i-th document. To find the closest emotion type of the clause, we calculate the Euclidean distance between $s^{i,j}$ and centers of the emotion domains, denoted as $\{sdc_1, sdc_2, ..., sdc_7\}$, then select the minimum distance $dis^{i,j}$. For each emotion domain, the domain radius is denoted as rad_i, which is calculated by Formula (3):

$$rad_i = \frac{1}{|sd_i|}\sum_{j=1}^{|sd_i|}\text{Euclidean}(sdc_i, v_j) \tag{3}$$

$$dis^{i,j} = \min(\text{Euclidean}(sdc_k, s^{i,j})) \tag{4}$$

where $|sd_i|$ is the number of emotion clauses in the emotion domain and v_j represents each clause in the emotion domain. If $dis^{i,j}$ is less than or equal to the corresponding minimum distance, we consider that the clause is more likely to be an emotion clause. Sentiment score $simi^{i,j} \in (0, 1]$ is calculated for screening EC-Pairs according to the distance between a clause and emotion domains.

$$simi^{i,j} = \frac{1}{1+dis^{i,j}} \tag{5}$$

In order to minimize the propagation of error in the classification-pairing mode, we do not directly classify the output vectors of BiLSTM. Each clause is regarded as both cause and emotion, and candidate EC-Pairs are constructed by cross-combination, organized as a two-dimensional matrix. We concatenate the vectors of the clauses in candidate EC-Pairs, then pass them into a linear layer to calculate the probability, denoted as $pro^{i,j}$.

$$pro^{i,j} = \text{Linear}([r^e; r^c]) \tag{6}$$

In Formula (6), r^e and r^c are the output features encoded by the clause-level BiLSTM, and $[\cdot;\cdot]$ means concatenation.

As shown in Figure 4, the red grid in the figure represents the ground truth EC-Pair, and the maximum number of clauses in this batch of documents is denoted as *m*. If the number of clauses is less than m, it needs to be aligned by padding zero. To obtain the probability matrix with the shape of (bs, m, m) by 3D convolution operation., we padded zero in the shadow grids at the boundary in the figure.

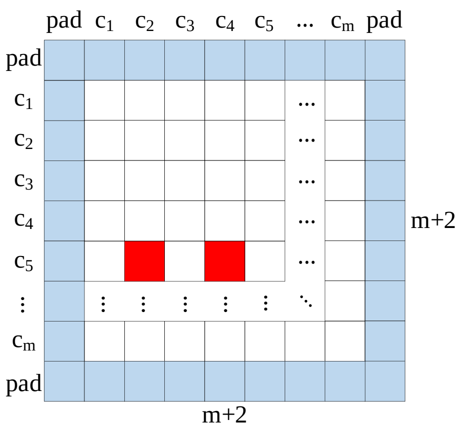

Fig. 4. Candidate EC-Pair in a document

3.5. 3D Fusion Extraction

We analyzed the original corpus of the ECPE dataset and found that the source of the corpus is the news reports of SINA NEWS‡ in three years, and documents are similar in structure. Considering the normalization of news report writing, we hypothesize that clauses in different documents have strong similarities in position. To capture this similar feature, we stack multiple documents into 3D shapes according to the structure of Figure 4, then pass them into the 3D-CNN network with a 3D convolution kernel.

Different from Figure 4, each element in the 3D matrix $element^{i,j}$ is composed of two dimensions, i.e., the channel of 3D-CNN is 2:

$$element^{i,j} = [\alpha simi^{i,j}, \beta pro^{i,j}] \tag{7}$$

where α, β are the weight parameters of the output of the two sub-modules, which are obtained by experiments. After performing the 3D convolution operation, the model learns similar features between adjacent documents and obtains the probability matrix $\mathbf{W}^p \in \mathbb{R}^{bs\times m\times m}$ of all candidate EC-Pairs through a Softmax function.

$$\mathbf{W}^p = \mathrm{Softmax}(3\mathrm{D\text{-}CNN}(\mathbf{W})) \tag{8}$$

where $\mathbf{W} \in \mathbb{R}^{bs\times(m+2)\times(m+2)}$ is the input matrix of 3D-CNN. To keep the stability of the sentiment vector generated by the pre-trained sentiment analysis model, parameters of the sentiment analysis module are frozen during training. The loss of 3D-CNN is obtained by calculating the output of 3D-CNN with cross-entropy loss function:

$$Loss = -\frac{1}{N}\sum_{\hat{y}^{i,j}\in\mathbf{W}^p}[y^{i,j}\log(\hat{y}^{i,j}) + (1-y^{i,j})\log(1-\hat{y}^{i,j})] \tag{9}$$

where N is the number of EC-Pairs, $\hat{y}^{i,j}$ and $y^{i,j}$ are the predicted label and ground-truth label of $element^{i,j}$.

4. Experiments

4.1. Dataset & Metrics

In 29, Gui et al. annotated 2105 documents from 20000 articles on the SINA NEWS website for the ECE task. However, documents containing multiple EC-Pairs are split into multiple samples, which is inconsistent with the actual scenario. Study 2 combined and annotated the ECE benchmark dataset, then constructed the ECPE benchmark dataset. The data distribution is shown in Table 2.

Each document in the dataset contains several clauses constituting one or more EC-Pairs. Compared with the ECE benchmark dataset, about 10% of the documents contain multiple EC-Pairs. In order to extract all EC-Pairs, we enumerate all possible clause

‡ https://news.sina.com.cn/

combinations in a document, then pass the 3D matrix composed of multiple documents into the 3D-CNN network to calculate the EC-Pair probability matrix $\mathbf{W}^p$ for each document.

Table 2. Distribution of ECPE dataset

	Number	Percentage
Number of EC-Pair in Document equals 1	1746	89.77%
Number of EC-Pair in Document equals 2	177	9.10%
Number of EC-Pair in Document greater than 2	22	1.13%
All	1945	100%

Figure 5(a) shows the five clauses contained in the document {C1, C2, C3, C4, C5}, where C5 is the emotion clause (the type of emotion word here is Happiness); Figure 5(b) describes the causal relationship between the clauses in this document. C5 caused by both C2 and C4, i.e., the EC-Pairs contained in the document are {C5←C2} and {C5←C4}.

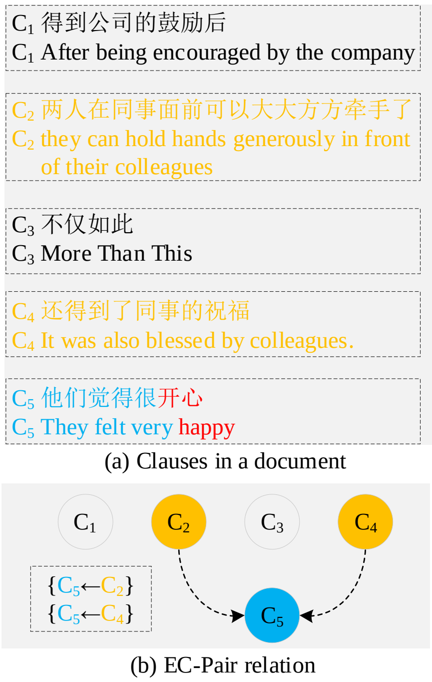

Fig. 5. Example in dataset

We construct two datasets for training and testing, the ratio of them is 9:1, then calculate the average score using ten-fold cross-validation. The same as the baseline model, we apply P(Precession), R(Recall), and F1 score as the metrics of our model.

4.2. Experimental Settings

Compared with word2vec 33 method, BERT 34 is more robust. The Chinese fine-tuning BERT[§] model generates word vectors of clauses, and the dimension is 768. In the encoding process, the number of hidden layer units of BiLSTM equals 100. In the 3D fusion extraction module, the convolution kernel size of 3D-CNN is set to (3, 3, 3). We apply dropout 35 to the linear layer to avoid overfitting and set the dropout rate to 0.3. The default weight and bias of BiLSTM, 3D-CNN, and linear layer are initialized by the Xavier method 36.

During the training period, to make the model's convergence process smoother, the learning rate is set to 3e-5, and introduced Adam 37 as the optimizer of the model. In addition, the batch size is 16 and the epoch of training is 100.

4.3. Compared ECPE Models

We compare these ECPE models with SEE-3D:

Baselines: This model is the benchmark model when Study 2 first proposed the ECPE task, including three baseline models: Indep, Inter-CE, and Inter-EC. This model divides the extraction into two steps:(1) Classifying each clause into emotion clause or cause clause to combine candidate EC-Pairs; (2) Eliminating EC-Pairs of low probability.

RANKCP: This extraction method is proposed by Study 24. The author used a graph attention mechanism to learn the interaction between emotion clauses and cause clauses in documents for filtering EC-Pairs.

E2EECPE: This end-to-end framework was proposed by Study 13. Biaffine attention is utilized to predict the causal relationship between clauses, Song transforming the clause pairing problem into the edge prediction problem.

ECPE-2D: This model is an improved method for **ECPE-2Steps** by Study 25. They represent EC-Pairs as a 2D structure that models the interactions between different EC-Pairs through two different 2D Transformers (Windows-constrained, Cross-road).

DQAN: The author used a dual-questioning attention mechanism to obtain the interaction between clauses in EC-Pair and other clauses, which improved the semantic understanding ability of the model 27.

Trans-ECPE: In 23, Fan et al. regard the EC-Pair extraction process as the construction process of a directed graph. The model regards clauses as nodes of the graph and extracts EC-Pairs while constructing the graph. The two versions of LSTM and BERT represent the different structures used in the model coding characteristics.

[§] https://huggingface.co/bert-base-chinese

4.4. Experimental Discussion

The SEE-3D model and other models are verified on the ECPE benchmark dataset, Table 3 illustrates the results of experiments. In addition to the EC-Pair extraction results, the emotional extraction and cause extraction results of models are also listed in the table.

(1) Our model achieved the **highest precession and F1 score** in **emotion** extraction, 88.43%, and 86.74%. In these comparative models, **Indep** extracts the causes and results separately, without fully considering the interaction between clauses. **Inter-CE** and **Inter-EC** add emotion(or cause) clauses to the classification process of cause(or emotion) clauses, increasing the accuracy by about 1%-2%. Inspired by these, the subsequent models are constructed to explore the internal relationship between different clauses. **RANKCP** constructs global features between clauses and finally achieves the highest recall rate of 87.03% when extracting emotion clauses. Based on this skill, **RANKCP** achieving the highest recall rate of 66.98% when extracting EC-Pairs. The sentiment analysis module of **SEE-3D** calculates the emotional similarity and forward feedbacks the emotional features of clauses in EC-Pairs, which helps to better identify emotion clauses.

(2) In the **EC-Pair** extraction, the precession and the F1 score were finally improved. Compared with the LSTM model, the BERT model used in **Trans-ECPE** coding has significantly improved in all aspects. Besides, **Trans-ECPE** based on BERT has achieved the highest accuracy and F1 score in cause extraction, 75.62%, and 69.74%. **ECPE-2D** based on Inter-EC is selected for comparison because it integrates emotional features into the cause clauses classification process, similar to **SEE-3D**. So we can better mine the influence of sentimental intensity on the extraction results. with **Trans-ECPE**(BERT), the extraction accuracy of emotion clauses increases by 1.27%, and the F1 score increases by 2%, which further improves the performance of 3D fusion extraction.

Table 3. Comparison results with other ECPE models. '♠' denotes the model in this paper. '-' denotes the original paper that had not reported the evaluation result

Model	Emotion Extraction			Cause Extraction			EC-Pair Extraction		
	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)
Baselines (Indep)	83.75	80.71	82.10	69.02	56.73	62.05	68.32	50.82	58.18
Baselines (Inter-CE)	84.94	81.22	83.00	68.09	56.34	61.51	69.02	51.35	59.01
Baselines (Inter-EC)	83.64	81.07	82.30	70.41	60.83	65.07	67.21	57.05	61.28
RANKCP	85.48	**87.03**	84.06	68.24	69.27	67.43	66.10	**66.98**	65.46
E2EECPE	85.95	79.15	82.38	70.62	60.30	65.03	64.78	61.05	62.80
ECPE-2D (Inter-EC)	85.37	81.97	83.54	71.51	62.74	66.76	71.73	57.54	63.66
DQAN							67.33	60.40	63.62
Trans-ECPE (LSTM)	80.80	84.39	82.56	67.42	65.34	66.36	65.15	63.54	64.34
Trans-ECPE (BERT)	87.16	82.44	84.74	**75.62**	**64.71**	**69.74**	**73.74**	63.07	67.99
SEE-3D♠	**88.43**	85.12	**86.74**	70.52	63.67	66.92	72.58	66.32	**69.31**

Overall, **SEE-3D** improves the extraction performance by deepening the understanding of emotional semantic information. The above analysis shows that the emotional dependence between clauses has important research value in the ECPE task.

4.5. Ablation Study

To investigate the effectiveness of sentiment analysis, we removed the calculation of emotion domain and emotional similarity when filtering EC-Pairs, i.e., the model only retains the clause encoder, linear layer, and 3D-CNN. When the model only retains the output value of the linear layer as the probability, it is difficult for 3D-CNN to learn the sentimental intensity interaction between documents. The extraction results of EC-Pairs are shown in Table 4. The F1 score of cause extraction is almost unchanged, but the precession, recall, and F1 score of emotion extraction are reduced by 3.82%, 4.23%, and 4.03%. Although the similarity of some cause clauses is learned in 3D-CNN, due to the lack of emotional semantic information, the final F1 score of EC-Pair extraction is 6.02% lower than that of the original **SEE-3D** model.

From the discussion above, we can find that emotional semantic information significantly influences the ECPE task. Due to the starting point of the ECPE task serving for the emotional reasoning task, the fusion of the information learned by sentiment analysis and cause extraction can improve the performance of our model.

Table 4. Ablation Study of Sentiment Analysis. We removed the emotional similarity metric from our model to reveal the effect of the sentiment analysis sub-module, denoted as "-SA"

Model	Emotion Extraction			Cause Extraction			EC-Pair Extraction		
	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)	P(%)	R(%)	F1(%)
SEE-3D (-SA)	84.61	80.89	82.71	70.39	**63.75**	66.90	65.46	61.25	63.29
SEE-3D	**88.43**	**85.12**	**86.74**	**70.52**	63.67	**66.92**	**72.58**	**66.32**	**69.31**

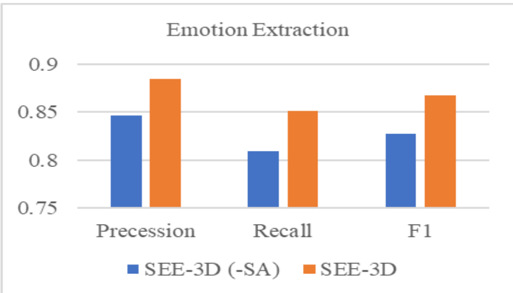

Fig. 6. Ablation Study of Sentiment Analysis

By analyzing the corpus source of the ECPE benchmark dataset, we proposed the introduction of 3D-CNN learning similar features of samples. To prove the effectiveness of 3D operation, we removed 3D-CNN from SEE-3D, and the comparison results are listed in Table 5. When our model directly uses the weighted sum of emotional similarity and linear layer output as the prediction results, the precession, recall, and F1 score are reduced by 2.39%, 2.17%, and 2.28%. In short, 3D-CNN learned the similarity between samples.

Table 5. Ablation Study of 3D-CNN. We removed the 3D-CNN from our model and directly used elements in Wp as extraction results, denoted as "-3D Conv"

Model	EC-Pair Extraction		
	P(%)	R(%)	F1(%)
SEE-3D (-3D Conv)	70.19	64.15	67.03
SEE-3D	**72.58**	**66.32**	**69.31**

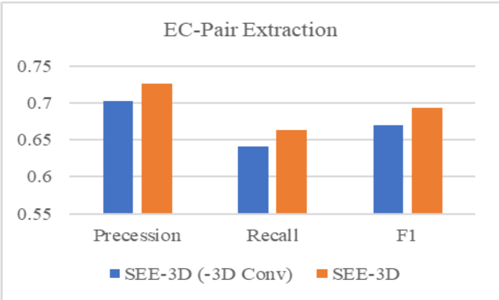

Fig. 7. Ablation Study of 3D-CNN

5. Conclusions

As an essential task in cause detection, ECPE can be applied to emotional reasoning in daily life and is helpful to improve the service quality of psychological counseling. To make emotion-cause detection more efficient, a sentiment-driven 3D extraction model SEE-3D is proposed. Our model integrates sentiment analysis and 3D convolution into the EC-Pairs extraction process and improves extraction performance. In this paper, the achievements can be summarized as follows:

(1) A pre-trained sentiment analysis model is integrated into the ECPE model effectively and enhances the extraction performance. We use the k-means clustering algorithm to process and analysis the emotional clauses, making the model further obtain the emotion domain distribution. The distance between the clause and emotion domain is used to represent the sentimental intensity of the clause.

(2) A 3D fusion extraction method is proposed to learn similar features of documents, which improves the accuracy of ECPE. The outputs of sentiment analysis are combined with the classical pairing algorithm. Thus, the model can learn the similar characteristics between samples, which effectively improves the accuracy of extraction.

The experiments on the ECPE benchmark dataset show that our model can learn the interaction between emotion and cause in documents.

Similar to other methods, our extraction method also has its limitations. It has a remarkable effect only when the samples are relatively uniform and the text writing format has the characteristics of a template. In other cause detection datasets, the 3D fusion extraction method does not necessarily have generalization ability. To further study datasets in other fields, we will continue to conduct research on cause detection and try to apply it to business scenarios.

Acknowledgments. This work was supported by the Graduate Innovation Foundation Project of Anhui University of Science & Technology: [grant number 2021CX2112], the National Natural Science Foundation of China (Grant NO.62076006), the University Synergy Innovation Program of Anhui Province (GXXT-2021-008), and the Anhui Provincial Key R&D Program (202004b11020029).

References

1. Xu X, Wu, H and Zhu, G. SCA-ECPE: Emotion-Cause Pair Extraction Based on Sentiment Clustering Analysis. In 2022 International Conference on Applications and Techniques in Cyber Intelligence (ATCI). (2022)
2. Xia, R. and Ding, Z., Emotion-Cause Pair Extraction: A New Task to Emotion Analysis in Texts. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 1003-1012. (2019)
3. Zhang, S., Hu, Z., Zhu, G., Jin, M., & Li, K. C. Sentiment classification model for Chinese micro-blog comments based on key sentences extraction. Soft Computing, Vol. 25, No. 1, 463-476. (2021)
4. Zhang, S., Yu, H., & Zhu, G. An emotional classification method of Chinese short comment text based on ELECTRA. Connection Science, Vol. 34, No. 1, 254-273. (2022)
5. Xu, H., Zhang, S., Zhu, G., & Zhu, H. ALSEE: a framework for attribute-level sentiment element extraction towards product reviews. Connection Science, Vol. 34, No. 1, 205-223. (2022)
6. Lee, S. Y. M., Chen, Y., & Huang, C. R. A text-driven rule-based system for emotion cause detection. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text, Los Angeles, CA, 45-53. (2010)
7. Xia, R., Zhang, M., & Ding, Z. RTHN: A rnn-transformer hierarchical network for emotion cause extraction. arXiv preprint arXiv:1906.01236. (2019)
8. Ding, Z., He, H., Zhang, M., & Xia, R. From independent prediction to reordered prediction: Integrating relative position and global label information to emotion cause identification. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, Hawaii, USA, Vol. 33, No. 1, 6343-6350. (2019)
9. Diao, Y., Lin, H., Yang, L., Fan, X., Chu, Y., Wu, D., & Xu, K. Emotion cause detection with enhanced-representation attention convolutional-context network. Soft Computing, Vol. 25, No. 2, 1297-1307. (2021)
10. Hu, G., Lu, G., & Zhao, Y. FSS-GCN: A graph convolutional networks with fusion of semantic and structure for emotion cause analysis. Knowledge-Based Systems, 212, 106584. (2021)
11. Yu, W., & Shi, C. Emotion Cause Extraction by Combining Intra-clause Sentiment-enhanced Attention and Inter-clause Consistency Interaction. In 2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS), 146-150. IEEE. (2021)
12. Yan, H., Gui, L., Pergola, G., & He, Y. Position Bias Mitigation: A Knowledge-Aware Graph Model for Emotion Cause Extraction. In Proceedings of the 59th Annual Meeting of

the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), 3364-3375. (2021)
13. Song, H., Zhang, C., Li, Q., & Song, D. End-to-end emotion-cause pair extraction via learning to link. arXiv preprint arXiv:2002.10710. (2020)
14. Fan, R., Wang, Y., & He, T. An end-to-end multi-task learning network with scope controller for emotion-cause pair extraction. In CCF International Conference on Natural Language Processing and Chinese Computing, 764-776. Springer, Cham. (2020)
15. Tang, H., Ji, D., & Zhou, Q. Joint multi-level attentional model for emotion detection and emotion-cause pair extraction. Neurocomputing, 409, 329-340. (2020)
16. Wu, S., Chen, F., Wu, F., Huang, Y., & Li, X. A multi-task learning neural network for emotion-cause pair extraction. In ECAI 2020, 2212-2219. IOS Press. (2020)
17. Chen, Y., Hou, W., Li, S., Wu, C., & Zhang, X. End-to-end emotion-cause pair extraction with graph convolutional network. In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain, 198-207. (2020)
18. Chen, X., Li, Q., & Wang, J. A unified sequence labeling model for emotion cause pair extraction. In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain, 208-218. (2020)
19. Cheng, Z., Jiang, Z., Yin, Y., Yu, H., & Gu, Q. A symmetric local search network for emotion-cause pair extraction. In Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain, 139-149. (2020)
20. Ding, Z., Xia, R., & Yu, J. End-to-end emotion-cause pair extraction based on sliding window multi-label learning. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 3574-3583. (2020)
21. Yuan, C., Fan, C., Bao, J., & Xu, R. Emotion-cause pair extraction as sequence labeling based on a novel tagging scheme. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 3568-3573. (2020)
22. Fan, C., Yuan, C., Gui, L., Zhang, Y., & Xu, R. Multi-task sequence tagging for emotion-cause pair extraction via tag distribution refinement. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29, 2339-2350. (2021)
23. Fan, C., Yuan, C., Du, J., Gui, L., Yang, M., & Xu, R. Transition-based Directed Graph Construction for Emotion-Cause Pair Extraction. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 3707-3717. (2020)
24. Wei, P., Zhao, J., & Mao, W. Effective Inter-Clause Modeling for End-to-End Emotion-Cause Pair Extraction. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 3171-3181. (2020)
25. Ding, Z., Xia, R., & Yu, J. ECPE-2D: emotion-cause pair extraction based on joint two-dimensional representation, interaction and prediction. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 3161-3170. (2020)
26. Singh, A., Hingane, S., Wani, S., & Modi, A. An End-to-End Network for Emotion-Cause Pair Extraction. In Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Zhengzhou, China, Sentiment and Social Media Analysis, 84-91. (2021)
27. Sun, Q., Yin, Y., & Yu, H. (2021, July). A Dual-Questioning Attention Network for Emotion-Cause Pair Extraction with Context Awareness. In 2021 International Joint Conference on Neural Networks (IJCNN), 1-8. IEEE. (2021)
28. Yu, J., Liu, W., He, Y., & Zhang, C. A mutually auxiliary multitask model with self-distillation for emotion-cause pair extraction. IEEE Access, 9, 26811-26821. (2021)
29. Gui, L., Wu, D., Xu, R., Lu, Q., & Zhou, Y. Event-Driven Emotion Cause Extraction with Corpus Construction. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, Austin, Texas, USA, 1639-1649. (2016)
30. Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. An efficient k-means clustering algorithm: Analysis and implementation. IEEE transactions on pattern analysis and machine intelligence, Vol. 24, No. 7, 881-892. (2002)

31. Graves, A., Mohamed, A. R., & Hinton, G. Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing, 6645-6649. IEEE. (2013)
32. Bahdanau, D., Cho, K., & Bengio, Y. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473. (2014)
33. Mikolov, T., Chen, K., Corrado, G., & Dean, J. Efficient Estimation of Word Representations in Vector Space. arXiv preprint arXiv:1301.3781. (2013)
34. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. (2018)
35. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, Vol. 15, No. 1, 1929-1958. (2014)
36. Glorot, X., & Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings, Sardinia, Italy, 249-256. (2010)
37. Kingma, D. P., & Ba, J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. (2014)

Xin Xu, born in 1991, Master candidate at Anhui University of Science and Technology. His main research interests include relation extraction and machine learning.

Guangli Zhu, born in 1969. Master, Associate professor, Master supervisor. Her current research interests include Web Mining, Semantic Search, and Calculation theory.

Houyue Wu, born in 1996. Master candidate at Anhui University of Science and Technology. His main research interests are Adversarial Sample Generation and Relation Extraction.

Shunxiang Zhang, born in 1970. PhD, professor, PhD supervisor. He is an professor at Anhui University of Science and Technology, China. His current research interests include Web Mining, Semantic Search, and Complex network.

KuanChing Li, born in 1967, professor, PhD supervisor. He is a University Distinguished Professor at Providence University, Taiwan. His current research interests include Parallel and Distributed Computing, Big Data and Emerging Technologies.

Received: March 03, 2022; Accepted: September 20, 2022.