24th European Conference on Artificial Intelligence - ECAI 2020
Santiago de Compostela, Spain

# A Multi-Task Learning Neural Network for Emotion-Cause Pair Extraction

**Sixing Wu**[1] and **Fang Chen**[2] and **Fangzhao Wu**[3] and **Yongfeng Huang**[4] and **Xing Li**[5]

**Abstract.** Emotion-cause pair extraction, which aims at extracting both the emotion and its corresponding cause in text, is a significant and challenging task in emotion analysis. Previous work formulated the task in a two-step framework, i.e., emotion and cause extraction, and emotion-cause relation classification. However, different tasks may correlate with each other and the two-step framework does not fully exploit the interactions between tasks. In this paper, we propose a multi-task neural network to perform emotion-cause pair extraction in a unified model. The task of relation classification is learned together with emotion and cause extraction. To this end, we develop a method to obtain training samples for relation classification without the dependence on the result of emotion and cause extraction. To fully exploit the interactions between different tasks, our model shares useful features across tasks. Moreover, we propose a method to incorporate position-aware emotion information in cause extraction to further improve the performance. Experimental results show that our model outperforms the state-of-the-art model on emotion-cause pair extraction.

## 1 INTRODUCTION

Emotion is a significant type of information in natural language, and there have been extensive research efforts devoted to detecting emotions expressed in texts. While the recognition of *emotion category* (e.g., happiness, fear, anger, etc.) has been widely studied [1, 21, 25, 22], the extraction of *emotion cause*, which aims at inferring the potential cause of an expressed emotion, receives more and more attention in recent years [17, 4, 10]. Because the emotion is usually triggered by some internal or external events [26, 24, 19], emotion cause extraction (ECE) takes an important role in extracting more comprehensive emotion information and has a wide range of application scenarios.

Lee et al. [15] presented an early definition of emotion cause in text, which refers to the explicit expressions (arguments or events) closely related to a certain emotions. In previous research, the emotion cause is annotated in different units and formulated as various problems [15, 2, 23, 25, 9]. Recent studies mostly formulated the extraction task as a clause-level classification problem [10, 27] to identify the clause which mentions the cause corresponding to a specific emotion [4, 28, 18, 9].

Most previous works on ECE supposed the emotion label is provided and perform cause extraction based on the annotated emotion. The requirement of emotion label limits the application of cause extraction and does not fully exploit the relationship between emotion and cause expressions [27]. To overcome the limitations, Xia and Ding [27] defined the task of emotion-cause pair extraction (ECPE), in which neither the emotion clause nor the cause clause is annotated. Given a document, the objective of ECPE is to obtain pairs of emotions and their corresponding causes at clause level, as illustrated in Figure 1. Compared to ECE, ECPE is a more challenging problem because of the unavailability of emotion information.

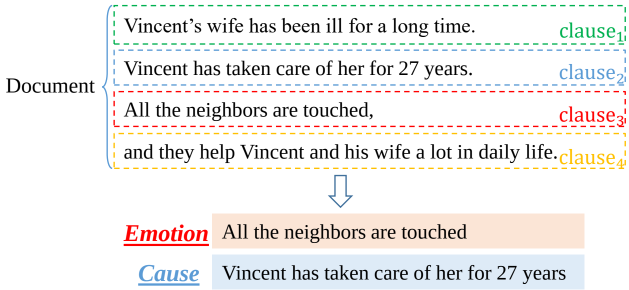



**Figure 1.** Illustration of the task of emotion-cause pair extraction (ECPE).

Intuitively, ECPE is performed via solving three individual tasks, i.e., emotion extraction, cause extraction, and identification of relevant emotion and cause. Based on this intuition, previous work formulated ECPE as a two-step framework [27]. Emotions and causes are extracted, and then related emotions and causes are paired.

Although proved to be effective, the previous two-step framework suffers from two shortcomings. Firstly, the relations between emotions and causes are classified after they have been extracted. However, the relations may contain useful semantic information for emotion and cause extraction. If it is highly probable that two clauses are in a causal relation and one of them contains emotional expressions, then the other is very likely to mention the cause. The two-step method does not fully exploit the relation information. Secondly, the errors made in the first step will be propagated to the second step. In the previous method, the clause representations obtained in the first step are taken as the input for the relation classification model. If the first model generates inaccurate clause representations, the performance of the second model will be negatively influenced.

In this paper, we propose a multi-task learning neural network to perform ECPE in a unified model (MTNECP). The task of relation classification is learned jointly with emotion and cause extraction, which means the input of relation prediction has to be independent of results of emotion and cause extraction. To this end, we develop

[1] Tsinghua University, China, email: wu-sx15@mails.tsinghua.edu.cn
[2] Tsinghua University, China, email: f-chen17@mails.tsinghua.edu.cn
[3] Microsoft Research Asia, China, email: wufangzhao@gmail.com
[4] Tsinghua University, China, email: yfhuang@tsinghua.edu.cn
[5] Tsinghua University, China, email: xing@cernet.edu.cn

a method to construct the training set for relation classification. To model the interactions between tasks, some useful features are shared to exploit correlations across tasks. Features for relation classification are shared with emotion and cause extraction to exploit relation-emotion and relation-cause interactions. Furthermore, we propose a method to incorporate the position-aware emotion information in cause extraction. To evaluate the effectiveness of our model, we conduct experiments on a publicly available dataset for ECPE.

The main contributions of this work can be summarized as follows:

- We propose a multi-task learning neural network to handle ECPE in a unified model, which learns emotion-cause relation classification jointly with emotion extraction and cause extraction, and shares useful features across tasks.
- We propose a method to incorporate the position-aware emotion information, which enables the model to leverage contextual emotion information for cause extraction.
- We conduct experiments to demonstrate the effectiveness of our method. The experimental results show that our model significantly outperforms the baseline method in terms of F1-score [6].

# 2 RELATED WORK

## 2.1 Emotion cause extraction

Targeted at identifying potential causes of the expressed emotion, emotion cause extraction (ECE) has been extensively discussed and investigated in previous works, which can be categorized into rule-based methods [17, 15, 6, 16, 23] and machine learning based methods [2, 11, 3].

Rule-based methods focus on developing effective algorithms based on linguistic features and rules to capture cause events in different expressions. Lee et al. [17] constructed an emotion cause corpus and summarized seven groups of linguistic cues which could serve as indicators of cause events. Based on the summarized cues, Lee et al. [15] proposed a set of linguistic rules to extract emotion cause expressed in specific syntactic patterns. Gao et al. [6] combined cause events with fine-grained emotion classes and extracted causes based on the corresponding emotion and syntactic features.

Syntactic features have been proved effective in ECE not only in unsupervised rule-based methods, but also in machine learning based methods [2, 7]. Most machine learning based methods formulated the task as a clause-level classification task, i.e., to predict whether a clause mentions the cause of a certain emotion [10]. Chen et al. [2] formulated the task as a multi-label classification problem and conducted ECE based on pattern-based and semantic-based features. Gui et al. [11] took distance and POS pattern feature into consideration, and compared the sequence labeling model and the binary classification model. Gui et al. [10] proposed to extract emotion cause with multi-kernel SVMs based on the features in syntactic trees.

In recent years, there is an emerging trend to apply neural networks on the task of ECE. Most of the related studies focused on developing more effective models to extract features of clauses and generate clause representations for cause detection. Attention mechanism [5] has been widely used to capture more informative causal expressions. Gui et al. [9] formulated the problem as a question answering task, and provided answers (cause) of questions (emotion) based on a deep memory network combined with the convolution operation [13]. To model the correlation between emotion context and cause context interactively, Li et al. [18] proposed a co-attention neural network to solve the problem. Besides the semantic features, there are other information integrated in neural network models to further improve the performance. Xia et al. [28] and Ding et al. [4] incorporated the relative position and global label information into clause representations to make more accurate cause predictions.

[6] The code will be made publicly available at https://github.com/wusx00/MTNECP

## 2.2 Emotion-cause pair extraction

Although extensive efforts have been made to propose effective models for cause extraction, most of the previous works captured the cause based on the emotion texts, which means the emotion has to be annotated before cause extraction. Xia et al. [27] pointed out the limitations of previous research on ECE, and defined the task of emotion-cause pair extraction (ECPE). In their method, emotion-cause pairs are extracted in two separate steps. Candidate emotion clauses and cause clauses are extracted first, and a classification model is trained to identify emotion-cause pairs from all candidates.

Previous studies either extract cause individually [9], or extract emotion and cause clauses, and classify the relations between them subsequently [28]. Our work differs from the previous works in the integration of relation information. Our model learns the relation classification task jointly with emotion extraction and cause extraction, so that all the three tasks can be performed with a unified model. Furthermore, we prove the effectiveness of sharing relation information with both emotion and cause extraction to improve the model's performances on both tasks. In addition, we develop a method to incorporate position-aware emotion information in our model.

# 3 METHODOLOGY

## 3.1 Problem formulation

The objective of emotion-cause pair extraction (ECPE) is to extract pairs consisting of emotions and their corresponding causes from a given document [27]. Emotions and causes are extracted at the clause level. Each input of our model is a document split into multiple clauses, represented as $d = \{c_1, ..., c_n\}$, where $n$ is the number of clauses in the document. Each clause is a subsequence of the whole document, represented as $c_i = \{w_{i1}, ..., w_{im}\}$, where $m$ is the length of the clause. Following previous works, the extraction task is formulated into the clause-level classification problem.

The pair extraction task is formulated as three subtasks in our model, i.e., emotion extraction, cause extraction and emotion-cause relation classification. We propose a multi-task learning neural network to learn the three tasks in a unified model. Useful features are shared across tasks to model the interactions between different tasks, which are described in section 3.3. The predictions of the three tasks are utilized to obtain the final result of ECPE, which is described in section 3.4. Figure 2 illustrates the architecture of our model.

## 3.2 Clause encoder

Given the input document containing multiple clauses, our model obtains the representation for each clause with the clause encoder, which consists of four layers, i.e., embedding layer, word-level contextual layer, word-level attention layer and clause-level contextual layer. In order to capture a specific representation for each task, three clause encoders are constructed to generate the task-specific clause representation. Because clause encoders for different tasks have the

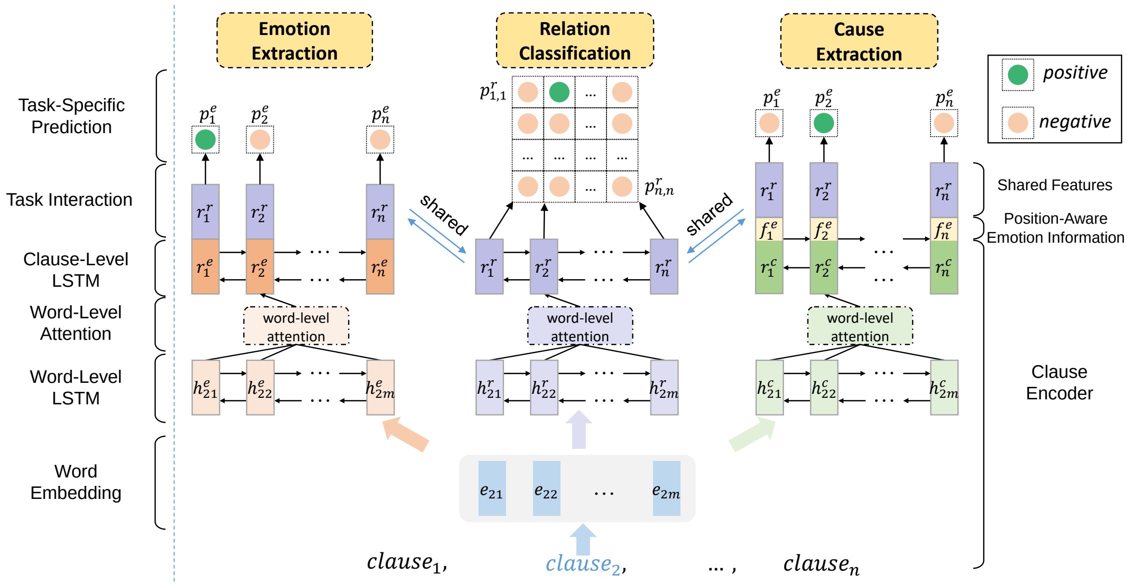



**Figure 2.** Illustration of the proposed multi-task learning neural network for emotion-cause pair extraction (MTNECP). For word embedding layer and word-level LSTM layer, the second clause is selected as the example.

same structure, we only describe the process of obtaining the clause representation for emotion extraction in the following.

The embedding layer transforms each word in the document to a $d_v$-dimensional vector. In this manner, the input word sequence is mapped to an embedded matrix. We denote $e_{c_i} = \{e_{i1}, e_{i2}, ..., e_{im}\}$ for the embedded matrix of the $i$-th clause in the document.

To encode the contextual information at word level, we employ a Bidirectional Long Short-Term Memory Network (BiLSTM) [12] to model the input sequence. The hidden state at each time-step is denoted as $h_{ij}^e = [\overrightarrow{LSTM}(e_{ij}); \overleftarrow{LSTM}(e_{ij})] \in \mathbb{R}^{2d}$, where $d$ is the number of hidden units in LSTM. We omit the details of BiLSTM because of limited space, which are presented in previous works [12, 8]. The word-level LSTM is performed on each clause to encode the inner-clause contexts.

Not all words are equally important in terms of expressing emotion information. An attention layer [20] is adopted to enable the model to focus on more informative words:

$$s_{ij}^e = u_e^T \tanh\left(W_a^e h_{ij}^e + b_a^e\right) \tag{1}$$

$$\alpha_{ij}^e = \frac{\exp\left(s_{ij}^e\right)}{\sum_{j=1}^{m} \exp\left(s_{ij}^e\right)} \tag{2}$$

where $W_a^e \in \mathbb{R}^{2d\times 2d}$ and $b_a^e \in \mathbb{R}^{2d}$ are weight matrix and the bias.

$h_i^e = \sum_{j=1}^{m} \alpha_{ij}^e h_{ij}^e$ denotes the representation of the $i$-th clause, which only encodes inner-clause contextual information. To encode the inter-clause contextual information, another BiLSTM layer is adopted at the clause level. $h_i^e$ is taken as the input, and we denote $r_i^e = [\overrightarrow{LSTM}(h_i^e); \overleftarrow{LSTM}(h_i^e)] \in \mathbb{R}^{2d}$ as the hidden state of the clause-level LSTM layer. $r_i^e$ is the representation of the $i$-th clause, which is utilized as the features for emotion extraction.

In the same way, our model encodes the contextual information for cause extraction and relation classification with two individual clause encoders, denoted as $r_i^c$ and $r_i^r$.

## 3.3 Multi-task learning framework

The clause encoders extract task-specific contextual features. However, it is not enough if the extracted features are utilized for each task independently, because there are interactions between different tasks, and features extracted for one task may provide useful information for another. To this end, features are shared from three aspects in our model to build up the relations between tasks, i.e., relation-emotion, relation-cause and emotion-cause.

### 3.3.1 *Emotion extraction*

The emotion extraction task is formulated as a clause-level classification problem. Features of each clause are computed based on the emotion-specific clause representation and the representation shared between relation and emotion:

$$z_i^e = [r_i^e; r_i^r] \tag{3}$$

$$p_i^e = \mathrm{softmax}(W_e z_i^e + b_e) \tag{4}$$

where $p_i^e$ is the probability distribution of emotion of the $i$-th clause. $W_e \in \mathbb{R}^{4d\times 2}$ and $b_e \in \mathbb{R}^2$ denote the weight matrix and the bias.

### 3.3.2 *Cause extraction*

Similar to the emotion extraction, the cause extraction task is performed through a clause-level classification. Besides the cause-specific clause representation and the representation shared between relation and cause, we incorporate position-aware emotion information as the additional feature, to model the interaction between emotion and cause extraction.

Relative position information has been proved effective in improving the performance on cause extraction [28, 4]. However, previous works incorporated relative information based on the annotated emotion clause. In ECPE, the emotion clause is unknown, thus the previous methods cannot be directly applied.

To solve the problem of the unavailability of emotion label, the probability distribution of emotion is leveraged as the emotion information without position information by previous work [27]. To incorporate position-aware emotion information, we apply a sliding window to acquire the emotion distribution of surrounding clauses for each clause. Specifically, for the $i$-th clause, the position-aware emotion information is encoded into a $2 * l_w + 1$ dimension vector, where $2 * l_w + 1$ is the length of the window:

$$f_i^e = [p_{i-l_w,1}^e; p_{i-l_w+1,1}^e; ...; p_{i+l_w,1}^e] \tag{5}$$

Note that the emotion extraction is formulated as a binary classification task (i.e., $p_i^e = \{p_{i,0}^e, p_{i,1}^e\}$), and only the probability of the positive class $p_{i,1}^e$ is taken, indicating whether there are emotion clauses close to the $i$-th clause. The motivation is based on the observation that short-term contexts usually provide useful causal information for emotion expression [2].

Different from the basic clause encoder proposed in section 3.2, the input of the clause-level LSTM layer for cause extraction is the concatenation of inner-clause representation and position-aware emotion feature vector. The hidden state of the modified clause-level LSTM layer is obtained based on the updated input, i.e., $r_i^c = [\overrightarrow{LSTM}([h_i^c; f_i^e]); \overleftarrow{LSTM}([h_i^c; f_i^e])]$.

The predicted probability for cause extraction $p_i^c$ is obtained in a similar way of emotion extraction. The clause-level contextual layer of relation classification is shared for cause extraction as well.

$$z_i^c = [r_i^c; r_i^r] \tag{6}$$

$$p_i^c = \text{softmax}(W_c z_i^c + b_c) \tag{7}$$

where $p_i^c$ is the probability distribution of the $i$-th clause. $W_c \in \mathbb{R}^{4d \times 2}$ and $b_c \in \mathbb{R}^2$ denote the weight matrix and the bias.

#### *3.3.3 Relation classification*

In ECPE, there could be multiple emotion and cause clauses in a document. Therefore, it is an important to determine whether an extracted cause is the reason for a specific emotion. In our model, we integrate the relation classification task into the unified model.

To predict whether the $j$-th is the clause serves as the cause of the $i$-th clause, our model makes the prediction based on the concatenation of representations of the two clauses.

$$z_{i,j}^r = [r_i^r; r_j^r] \tag{8}$$

$$p_{i,j}^r = \text{softmax}(W_r z_{i,j}^r + b_r) \tag{9}$$

where $p_{i,j}^r$ denotes the probability that the $j$-th clause is the cause of the $i$-th clause. $W_r \in \mathbb{R}^{4d \times 2}$ and $b_r \in \mathbb{R}^2$ denote the weight matrix and the bias.

Because our model learns relation classification jointly with emotion and cause extraction, unlike previous work which performed relation classification based on the result of emotion and cause extraction [27], the input of the relation classifier of our model has to be independent of the result of emotion and cause extraction.

An intuitive solution is to take all pairs of clauses in a document as the training samples. However, the dataset for relation classification will be extremely imbalanced in this way. To generate a better training set, we select a subset of all possible clause pairs.

Specifically, for each candidate pair, if the first clause of the pair is an emotion clause or the second clause of the pair is a cause clause, the pair will be selected as a training sample. In a word, during the training phase, the relation classifier is optimized only with pairs which contains the ground-truth emotion clause or cause clause or both of them. The effectiveness of the dataset construction method will be proved in section 4.4.

## 3.4 Emotion cause pair extraction algorithm

Our multi-task learning neural network solves three tasks in ECPE. The final emotion-cause pairs are extracted based on the results of the three tasks. Specifically, each pair of the extracted emotion clause $c_i$ and the cause clause $c_j$ is taken as a candidate pair. If the pair $(c_i, c_j)$ is predicted as positive by the relation classification layer, it is extracted as an emotion-cause pair. Because all the three tasks are performed in a unified model, the relation classification does not rely on the result of emotion and cause extraction.

Although our model is able to conduct ECPE in a unified model, we develop a modified two-step method to further improve the performance. In the unified model, The relation features are shared to the tasks of emotion and cause extraction. While sharing the features can improve model's performance on emotion and cause extraction, it may not achieve the best performance on relation classification. An additional individual relation classifier may further improve the final result of ECPE.

In the previous two-step method [27], the clause representations generated in the first model are taken as the input of the relation classifier, which may lead to the error propagation. Therefore, we learn an individual relation classifier with the word embedding as input, which has the same structure as the clause encoder proposed in section 3.2. The features are computed via concatenating three representations, i.e., representations of the pair of clauses, and the representation of the distance between the predicted clauses:

$$z_{i,j}^{r'} = [r_i^{r'}; r_j^{r'}; v_{i,j}^d] \tag{10}$$

$$p_{i,j}^{r'} = \text{softmax}(W_{r'} z_{i,j}^{r'} + b_{r'}) \tag{11}$$

where $p_{i,j}^{r'}$ denotes the probability that the $j$-th clause is the cause of the $i$-th clause. $W_{r'} \in \mathbb{R}^{(4d+d') \times 2}$ and $b_{r'} \in \mathbb{R}^2$ denote the weight matrix and the bias. $v_{i,j}^d$ is the representation of the distance between the two clauses, which is a $d'$-dimensional vector.

Similarly, the training set is obtained from all possible pairs based on the method proposed in section 3.3.3. We evaluate different relation classification methods in section 4.4.

## 3.5 Model training

The model is trained for multiple tasks jointly, and the global loss function is the weighted sum of cross-entropy loss of all tasks:

$$L = \lambda_e L_e + \lambda_c L_c + \lambda_r L_r + \lambda_{reg} ||\Theta||^2 \tag{12}$$

where $\lambda_e$, $\lambda_c$ and $\lambda_r$ are coefficients of the loss for all the tasks, and $\Theta$ denotes the parameter set for L2 regularization term.

## 4 EXPERIMENT

In this section, we present the evaluation of our proposed multi-task learning neural network for emotion-cause pair extraction.

**Table 1.** Performance comparison on the emotion-cause pair extraction dataset. The best results are highlighted in bold.

| | | Emotion Extraction | | | Cause Extraction | | | Emotion-Cause Pair Extraction | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model | ***P*** | ***R*** | ***F1*** | ***P*** | ***R*** | ***F1*** | ***P*** | ***R*** | ***F1*** |
| **Baselines** | Indep | 0.8375 | 0.8071 | 0.8210 | 0.6902 | 0.5673 | 0.6205 | 0.6832 | 0.5082 | 0.5818 |
| | InterCE | 0.8494 | 0.8122 | 0.8300 | 0.6809 | 0.5634 | 0.6151 | **0.6902** | 0.5135 | 0.5901 |
| | InterEC | 0.8364 | 0.8107 | 0.8230 | 0.7041 | 0.6083 | 0.6507 | 0.6721 | 0.5705 | 0.6128 |
| **Ablation Study** | MTNECP w/o Emotion | 0.8556 | **0.8487** | 0.8511 | 0.7269 | 0.6212 | 0.6675 | 0.6415 | 0.5778 | 0.6051 |
| | MTNECP w/o Relation | 0.8632 | 0.8319 | 0.8463 | 0.7268 | 0.6282 | 0.6719 | 0.6320 | 0.5904 | 0.6071 |
| | MTNECP w/o Filter | 0.8662 | 0.8393 | 0.8520 | 0.7400 | 0.6378 | 0.6844 | 0.6479 | **0.6046** | 0.6245 |
| **Full Model** | MTNECP | **0.8662** | 0.8393 | **0.8520** | **0.7400** | **0.6378** | **0.6844** | 0.6828 | 0.5894 | **0.6321** |

## 4.1 Experimental settings

### *4.1.1 Dataset*

In our experiment, we utilize the publicly available dataset for the task of ECPE constructed by Xia and Ding [27], which is annotated based on the benchmark dataset for the task of ECE [10]. The statistics of the dataset is presented in 2. Each input is a document split into multiple clauses, and the aim is to detect all pairs of emotions and causes at the clause level. Because most previous works focused on cause extraction with emotion annotation, we evaluate our model's performance on the ECE dataset [10] as well. Both datasets adopted the Ekmans research on emotion to determine the emotion of texts, which identifies six primary emotions, e.g., happiness, sadness, fear, anger, disgust and surprise. Clauses with primary emotions are labeled as emotion clauses in both datasets.

**Table 2.** Statistics of the ECPE dataset. Some clauses contain the emotion and cause at the same time.

| # of | Documents | # of | Clauses |
|---|---|---|---|
| All | 1945 | All | 28727 |
| 1 pair | 1765 | All emotion&cause | 3720 |
| 2 pairs | 159 | Emotion | 2085 |
| $\geq$ 3 pairs | 21 | Cause | 2142 |

### *4.1.2 Evaluation metrics*

Precision, recall and F1 are used as the evaluation metrics. Because the task aims at extracting emotion-cause pairs, the metrics are calculated based on the extracted pairs and ground-truth pairs, the same as [27]. In addition, we also evaluate the performance of emotion and cause extraction task via precision, recall and F1 based on the extracted emotion and cause clauses, as defined in [10].

### *4.1.3 Implementation details*

The model is implemented with Tensorflow[7] and trained on NVIDIA 1080Ti. The model is trained with Adam [14]. The initial learning rate is set to 0.005 and the batch size is set to 32. We randomly select 90% of the data as the training set and the rest 10% as test, following previous work [27]. 10-fold cross validation is used, and each process is repeated 2 times, with the average result as the final result.

The dimension of word vectors are set to 200, and initialized with the word embedding pre-trained on a Chinese Weibo corpora [28, 4, 27]. The numbers of hidden units in LSTM layers are set to 100. In the global loss function, $\lambda_e$, $\lambda_c$, and $\lambda_r$ are all set to 1, and $\lambda_{reg}$ is set to $10^{-5}$. $l_w$ is set to 3, which means the window size of the position-aware emotion information is set to 7.

[7] https://www.tensorflow.org

## 4.2 Compared methods

To validate the effectiveness of our model, we compare it with several baseline models.

- **Indep**: It is the two-step method [27]. Emotion and cause clauses are extracted based on a multi-task LSTM model. Then a relation classifier is learned to filter irrelevant emotion and cause clauses.
- **IndepCE / IndepEC**: They are two variants of **Indep** [27], which incorporate the predictions of cause (or emotion) extraction as features of emotion (or cause) extraction.
- **MTNECP**: It is our model which perform three tasks jointly.
- **MTNECP w/o Relation**: In this model, the impact of relation is removed to evaluate the effect of relation information.
- **MTNECP w/o Emotion**: In this model, the position-aware emotion information is removed to validate its effect.
- **MTNECP w/o Filter**: It is our model which does not remove irrelevant pairs based on the result of relation classification task.

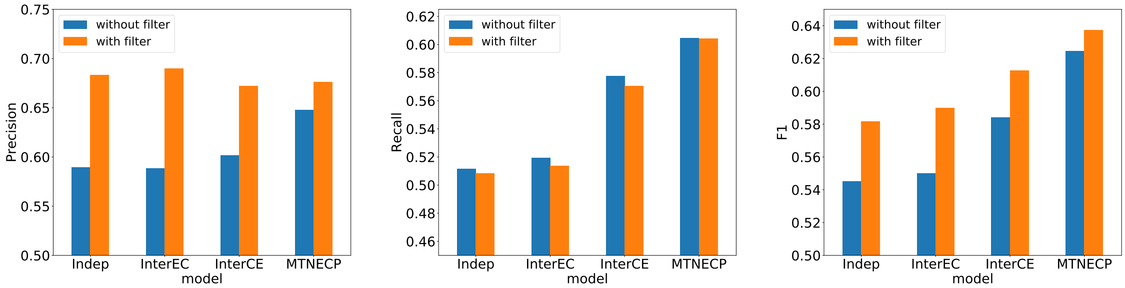
## 4.3 Main results

The experimental results of the compared models are presented in Table 1. It can be observed that our proposed MTNECP model achieves the best performance on ECPE in terms of F1. Compared with the best baseline model InterEC, our model gains obvious improvement on all evaluation metrics. The improvement on ECPE comes from both emotion extraction (2.90% by F1) and cause extraction (3.37% by F1), which demonstrates that the relation information can benefit both emotion and cause extraction, and our model effectively exploits interactions across tasks via multi-task learning and feature sharing.

We conduct ablation studies to investigate the contribution of different components of MTNECP. Firstly, based on the comparison between MTNECP and MTNECP w/o Emotion, we observe that the performance on ECPE drops significantly after removing the position-aware emotion information. Both precision and recall on cause extraction are negatively influenced, which proves that contextual emotion information is significant for cause extraction.

Secondly, via comparing the result of MTNECP and MTNECP w/o Filter, it can be demonstrated that in the document with multiple emotions and causes, the result of relation classification task effectively removes irrelevant pairs. While the recall drops 1.52%, the precision increases 3.49%, which brings an obvious improvement on F1. More comprehensive evaluations are reported in section 4.4.

Thirdly, the comparison between MTNECP w/o Filter and MTNECP w/o Relation validates the effectiveness of relation information. Both models have the same structure, but the impact of relation classification task $\lambda_r$ is set to zero in the objective of MTNECP




(a) Precision (b) Recall (c) F1

**Figure 3.** Influence of the individual relation classifier on the performance of different models.

w/o Relation. The result shows that both emotion and cause extraction gain improvements by integrating the relation classification task, which proves the positive impact of the relation information.

## 4.4 Results of relation classification models

The classification of relation is an important task in ECPE. To evaluate the effectiveness of different relation classification methods, we compare the results of several models, as presented in Table 3.

- **Indep/InterCE/InterEC w/o Filter**: They are the two-step methods without the pair filter, i.e., all pairs of extracted emotions and causes in a document are taken as emotion-cause pairs.
- **MTNECP + Filter/Filter(Encoder)**: They are the methods with an additional relation classifier as the pair filter. The former takes the clause representations of MTNECP as input features, and the latter adopts an individual clause encoder proposed in section 3.2.
- **MTNECP w/o Selection**: In MTNECP, we sample a subset of all clause pairs as the training set. In this model, no sampling is performed and all pairs of clauses are utilized during training.

**Table 3.** Effect of different relation classification methods. The best results are highlighted in bold.

| Model | *P* | *R* | *F1* |
|---|---|---|---|
| Indep w/o Filter | 0.5894 | 0.5114 | 0.5451 |
| Indep | 0.6832 | 0.5082 | 0.5818 |
| InterCE w/o Filter | 0.5883 | 0.5192 | 0.5500 |
| InterCE | 0.6902 | 0.5135 | 0.5901 |
| InterEC w/o Filter | 0.6019 | 0.5775 | 0.5842 |
| InterEC | 0.6721 | 0.5705 | 0.6128 |
| MTNECP w/o Selection | 0.6734 | 0.5916 | 0.6290 |
| MTNECP | 0.6828 | 0.5894 | 0.6321 |
| MTNECP w/o Filter | 0.6479 | **0.6046** | 0.6245 |
| MTNECP+Filter | 0.6763 | 0.6041 | 0.6373 |
| MTNECP+Filter(Encoder) | **0.6944** | 0.6017 | **0.6440** |

It can be observed from Figure 3 that filtering irrelevant emotions and clauses significantly improves the performance on ECPE, especially for methods which formulate the task in a two-step framework. Specifically, via adopting a relation classifier, the precision of ECPE is significantly improved in Indep, InterCE, InterEC. Meanwhile, the negative influence on recall is slight, which indicates that the classifier retains most of the correct pairs in candidates.

Compared with the baseline models, our model is able to achieve a higher performance in terms of F1 even without conducting pair filtering (MTNECP w/o Filter). Both the precision and recall are significantly improved. Because for all the baseline models, relation features are extracted only in the second step via training a relation classifier. As to our model, the relation features are extracted and shared with emotion and cause extraction, which provides beneficial information to both of the two tasks.

In MTNECP, the predictions of relation classification task are used to filter irrelevant emotions and causes. The comparison between MTNECP and MTNECP w/o Filter proves the effectiveness. However, the relation classifier in the unified model may not achieve the best performance on relation classification because its features are shared with the other two tasks. It can be observed that both MTNECP+Filter and MTNECP+Filter(Encoder) outperform MTNECP in terms of F1, which indicates that the performance on ECPE can be further improved via adopting an additional relation classifier.

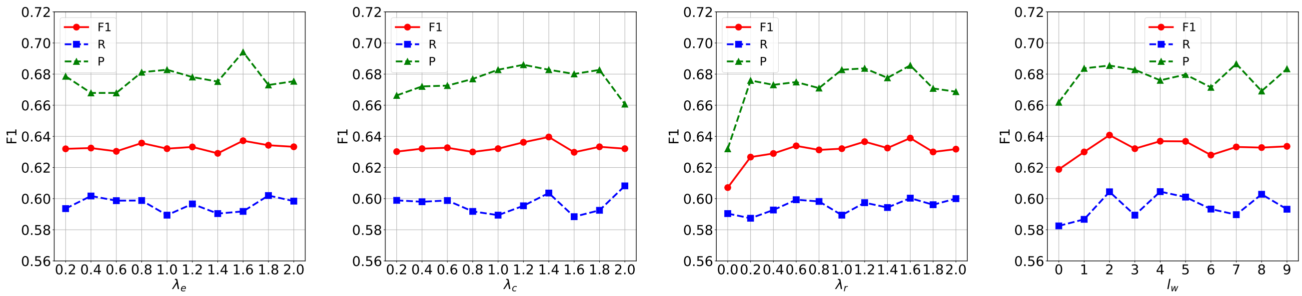
In addition, we make the comparison among two different relation classifiers. Compared with MTNECP+Filter, MTNECP+Filter(Encoder) gains an improvement of 1.81% in precision and 0.67% in F1. The result verifies our hypothesis that the error may be propagated if the clause representations of the first model are adopted as input features of the second, and an individual clause encoder is able to obtain more accurate clause representations for the relation classification task.

## 4.5 Hyper-parameter analysis

In this section, we present the evaluation of the impact of different hyper-parameters. The results are plotted in Figure 4.

- $\lambda_e$, $\lambda_c$ and $\lambda_r$ control the influence of emotion, cause extraction and relation classification respectively in the overall objective.
- $l_w$ is the radius of the sliding window which controls the number of surrounding clauses in the position-aware emotion information presented in section 3.3.2.

From the plot, we can observe that the performance in terms of F1 improves when $\lambda_r$ increases from 0 to 0.6. When $\lambda_r$ increases from 0.6 to 2, the performance fluctuates slightly and does not increase with $\lambda_r$. On the other hand, the overall performance on ECPE is not obviously influenced by the value of $\lambda_e$ and $\lambda_c$. The reason lies in the application of individual clause encoders for emotion and




(a) Influence of $\lambda_e$. (b) Influence of $\lambda_c$. (c) Influence of $\lambda_r$. (d) Influence of $l_w$.

**Figure 4.** The performance of our model with different settings of hyper-parameters.

cause extraction in MTNECP, which enables the model to extract task-specific contextual features.

As to the impact of $l_w$, it can be observed that the F1 value improves when $l_w$ increases from 0 to 2, which demonstrates that the emotion information benefits cause extraction and further improves the result of ECPE. When $l_w$ keeps increasing, the performance does not improve and fluctuates slightly, which indicates that long-term emotion information have little impact on cause extraction.

### 4.6 Case study

In order to make more explicit analysis of our model, we present some cases in Table 4. Because each document contains many clauses, only the emotion and cause clauses are listed in the table.

**Table 4.** Case study.

| Clauses | Label | MTNECP | InterEC |
|---|---|---|---|
| [It would hurt if you see that the little boy is so weak.]$_{c_1}$ | E:$c_1$ C:$c_1$ | E:$c_1$ C:$c_1$ | E:None C:None |
| [The medical record is put on the desk.]$_{c_1}$ [She fell into despair because of the death of her husband.]$_{c_2}$ | E:$c_2$ C:$c_2$ | E:$c_2$ C:$c_2$ | E:$c_2$ C:$c_1$ |
| [She was not happy.]$_{c_1}$ [Because she found that the type of her blood was O, and her mother's was AB and father's was A,]$_{c_2}$ [which means she was not the daughter of them.]$_{c_3}$ | E:$c_1$ C:$c_3$ | E:$c_1$ C:$c_2$ | E:$c_1$ C:$c_2$ |

As can be viewed, our model is able to identify emotion and cause clauses in more complicated cases. In the first example, the cause and emotion are mentioned in the same clause. In the second example, there is a clause describing an unrelated event to the emotion. Our model is able to recognize the cause correctly.

However, there are also some cases where both models fail to make accurate prediction. The causal relationship could be very implicit in natural language. There could be several events described in a same document, and only one of them is the real cause of the emotion. Therefore, further research needs to be conducted to develop more effective models to understand the implicit causal relation on emotions in texts in the future.

### 4.7 Evaluation on emotion cause extraction

As a related task of ECPE, ECE has been widely studied in previous works. To compare our model with a wider range of related works, we evaluate our model's performance on ECE, as reported in Table 5. The descriptions of baseline methods are omitted due to the limited space. All the baselines are mentioned and compared in previous works on ECE [28, 4, 27, 18], and the results of baselines are retrieved from previous works as well.

As can be observed, most previous methods depend on the annotation of emotions. The performance declines significantly while the emotion annotation is removed (CANN-E [28]), or the relative position to the emotion clause is unavailable (RTHN-APE [28]). Our model achieves the highest performance in terms of precision, recall and F1 compared with methods without reliance of the emotion annotation, and reaches comparable performance with many baseline methods dependent on emotion annotations.

**Table 5.** Performance comparison on the emotion cause extraction dataset. The best results are highlighted in bold.

| Model | *P* | *R* | *F1* | Emotion |
|---|---|---|---|---|
| RB [15] | 0.6747 | 0.4287 | 0.5243 | ✗ |
| CB [23] | 0.2672 | 0.7130 | 0.3887 | ✗ |
| RB + CB + ML [2] | 0.5921 | 0.5307 | 0.5597 | ✗ |
| Multi-kernel [10] | 0.6588 | 0.6927 | 0.6752 | ✗ |
| CNN [13] | 0.6215 | 0.5944 | 0.6076 | ✗ |
| ConvMS-MemNet [9] | 0.7076 | 0.6838 | 0.6955 | ✗ |
| CANN [18] | **0.7721** | 0.6891 | 0.7266 | ✗ |
| PAE-DGL [4] | 0.7619 | 0.6908 | 0.7242 | ✗ |
| RTHN [28] | 0.7697 | **0.7662** | **0.7677** | ✗ |
| RTHN-APE [28] | 0.5800 | 0.5618 | 0.5694 | ✓ |
| CANN-E [27] | 0.4826 | 0.3160 | 0.3797 | ✓ |
| Inter-EC [27] | 0.7041 | 0.6083 | 0.6507 | ✓ |
| MTNECP | **0.7400** | **0.6378** | **0.6844** | ✓ |

## 5 CONCLUSION

In this paper, we propose a unified multi-task neural network for the task of emotion-cause pair extraction. The model learns emotion extraction, cause extraction and relation classification jointly, and shares useful features to exploit interactions between tasks. In addition, we develop an individual relation classifier which can further improve the final performance on emotion-cause pair extraction. We demonstrate the effectiveness of our method via experiments on a benchmark dataset. The experimental results prove that our unified model significantly outperforms the baseline two-step model, and the performance on emotion and cause extraction can be obviously improved via integrating the relation information. Compared with baseline methods, our model is able to extract cause without reliance on the emotion annotation, and achieve the best performance in cases where no emotion annotation is available.

## 6 ACKNOWLEDGEMENT

We would like to thank the referees for their comments, which helped improve this paper considerably. This work is supported by the National Key Research and Development Program of China (No.2018YFC1604002) and National Natural Science Foundation of China (Grant nos. U1536201 and U1705261).

## REFERENCES

[1] Muhammad Abdul-Mageed and Lyle Ungar, 'Emonet: Fine-grained emotion detection with gated recurrent neural networks', in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 718–728, (2017).

[2] Ying Chen, Sophia Yat Mei Lee, Shoushan Li, and Chu-Ren Huang, 'Emotion cause detection with linguistic constructions', in *Proceedings of the 23rd International Conference on Computational Linguistics*, pp. 179–187. Association for Computational Linguistics, (2010).

[3] Xiyao Cheng, Ying Chen, Bixiao Cheng, Shoushan Li, and Guodong Zhou, 'An emotion cause corpus for chinese microblogs with multiple-user structures', *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, **17**(1), 6, (2017).

[4] Zixiang Ding, Huihui He, Mengran Zhang, and Rui Xia, 'From independent prediction to re-ordered prediction: Integrating relative position and global label information to emotion cause identification', in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, (2019).

[5] Bahdanau Dzmitry, Cho Kyunghyun, and Bengio Yoshua, 'Neural machine translation by jointly learning to align and translate', in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, (2015).

[6] Kai Gao, Hua Xu, and Jiushuo Wang, 'A rule-based approach to emotion cause detection for chinese micro-blogs', *Expert Systems with Applications*, **42**(9), 4517–4528, (2015).

[7] Diman Ghazi, Diana Inkpen, and Stan Szpakowicz, 'Detecting emotion stimuli in emotion-bearing sentences', in *International Conference on Intelligent Text Processing and Computational Linguistics*, pp. 152–165. Springer, (2015).

[8] Alex Graves, Abdel-rahman Mohamed, and Geoffrey Hinton, 'Speech recognition with deep recurrent neural networks', in *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 6645–6649. IEEE, (2013).

[9] Lin Gui, Jiannan Hu, Yulan He, Ruifeng Xu, Qin Lu, and Jiachen Du, 'A question answering approach to emotion cause extraction', in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 1593–1602. Association for Computational Linguistics, (2017).

[10] Lin Gui, Dongyin Wu, Ruifeng Xu, Qin Lu, and Yu Zhou, 'Event-driven emotion cause extraction with corpus construction.', in *EMNLP*, pp. 1639–1649. World Scientific, (2016).

[11] Lin Gui, Li Yuan, Ruifeng Xu, Bin Liu, Qin Lu, and Yu Zhou, 'Emotion cause detection with linguistic construction in chinese weibo text', in *Natural Language Processing and Chinese Computing*, 457–464, Springer, (2014).

[12] Sepp Hochreiter and Jürgen Schmidhuber, 'Long short-term memory', *Neural computation*, **9**(8), 1735–1780, (1997).

[13] Yoon Kim, 'Convolutional neural networks for sentence classification', in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1746–1751, (2014).

[14] Diederik P. Kingma and Jimmy Ba, 'Adam: A method for stochastic optimization', in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, (2015).

[15] Sophia Yat Mei Lee, Ying Chen, and Chu-Ren Huang, '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*, pp. 45–53. Association for Computational Linguistics, (2010).

[16] Sophia Yat Mei Lee, Ying Chen, Chu-Ren Huang, and Shoushan Li, 'Detecting emotion causes with a linguistic rule-based approach 1', *Computational Intelligence*, **29**(3), 390–416, (2013).

[17] Sophia Yat Mei Lee, Ying Chen, Shoushan Li, and Chu-Ren Huang, 'Emotion cause events: Corpus construction and analysis.', in *LREC*, (2010).

[18] Xiangju Li, Kaisong Song, Shi Feng, Daling Wang, and Yifei Zhang, 'A co-attention neural network model for emotion cause analysis with emotional context awareness', in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pp. 4752–4757, (2018).

[19] Bing Liu, *Sentiment analysis: Mining opinions, sentiments, and emotions*, Cambridge University Press, 2015.

[20] Thang Luong, Hieu Pham, and Christopher D. Manning, 'Effective approaches to attention-based neural machine translation', in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pp. 1412–1421, Lisbon, Portugal, (September 2015). Association for Computational Linguistics.

[21] Saif M Mohammad and Peter D Turney, 'Crowdsourcing a word–emotion association lexicon', *Computational Intelligence*, **29**(3), 436–465, (2013).

[22] Robert Plutchik, 'The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice', *American scientist*, **89**(4), 344–350, (2001).

[23] Irene Russo, Tommaso Caselli, Francesco Rubino, Ester Boldrini, and Patricio Martínez-Barco, 'Emocause: an easy-adaptable approach to emotion cause contexts', in *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, pp. 153–160. Association for Computational Linguistics, (2011).

[24] Daniel Schacter, Daniel Gilbert, Daniel Wegner, and Bruce M Hood, *Psychology: European Edition*, Macmillan International Higher Education, 2011.

[25] Yichen Wang and Aditya Pal, 'Detecting emotions in social media: A constrained optimization approach', in *Twenty-Fourth International Joint Conference on Artificial Intelligence*, (2015).

[26] Anna Wierzbicka, *Emotions across languages and cultures: Diversity and universals*, Cambridge University Press, 1999.

[27] Rui Xia and Zixiang Ding, '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*, p. 10031012, (2019).

[28] Rui Xia, Mengran Zhang, and Zixiang Ding, 'Rthn: A rnn-transformer hierarchical network for emotion cause extraction', in *the 28th International Joint Conference on Artificial Intelligence (IJCAI-2019)*, (2019).