Research Article

Siamese Interaction and Fine-Tuning Representation of Chinese Semantic Matching Algorithm Based on RoBERTa-wwm-ext

Baohua Qiang [ID],[1] Guangyong Xi [ID],[1] Yufeng Wang,[2] Xianyi Yang [ID],[1] and Yuemeng Wang[2]

[1]*Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China*
[2]*Hebei Key Laboratory of Intelligent Information Perception and Processing, The 54th Research Institute of CETC, Shijiazhuang 050081, China*

Correspondence should be addressed to Xianyi Yang; xianyiyang65@126.com

Semantic matching research is the cornerstone of research in the fields of natural language similarity measurement and sensor ontology matching (OM). In the existing Chinese semantic matching methods, there are some shortcomings, such as the single dimension of semantic expression, the insufficient expression of context semantic relations, and the insufficient interaction of semantic information between different sentences. This paper proposes a Chinese semantic matching algorithm based on RoBERTa-wwm-ext with Siamese interaction and fine-tuning representation (RSIFR). The RSIFR model initializes the model with RoBERTa-wwm-ext as a vector of text. Firstly, a Siamese structure with embedded soft alignment attention mechanism and BiLSTM is constructed to realize the information interaction between two sentences. Secondly, LSTM-BiLSTM network structure is constructed to enhance the expression of semantic logic before and after sentences. Then, build a training model with fine-tuning mechanism. Fine tune the text's eigenvector parameters through label supervision. Finally, the fusion vectors of the sentence pairs are inserted into the MLP network layer, resulting in semantic matching results. RSIFR model starts from a variety of dimensions, strengthens the expression ability of vectors to text semantic relations, deeply mines the semantic similarities and differences between different sentences, and generally improves the Chinese semantic matching performance. Experiments on the public dataset LCQMC show that our model outperforms existing Chinese semantic matching models.

1. Introduction

Chinese semantic matching is to judge the semantic match between two different texts for them. The core of Chinese semantic matching task lies in mining the deep semantic information of text and exploring the semantic relationship between different texts. The research of text semantic matching can be applied to application areas such as intelligent question and answer, machine translation [1], natural language inference [2], WEB Sensor Ontology Matching (OM), and Entity Semantic Similarity Measurement [3].

The text feature vector extracted by using deep neural network technology can improve the vector's ability to characterize the text semantics, but it also lacks the representation of the semantic relationship between two sentences in the Chinese semantic matching task. Niu et al. combined Siamese network structure with deep learning techniques such as BiLSTM to effectively extract deep features of the text [4]. Yang and Zhang applied the attention mechanism in the Chinese semantic matching task to improve the representation of text by feature vectors [5]. The feature vector of the text extracted by the Chinese preprocessing model related to BERT [6] has stronger semantic expression ability, which effectively improves the performance of the Chinese semantic matching model, but lacks the semantic information interaction between different texts.

In order to enhance the interactivity of semantic information and improve the ability of vectors to represent the semantic matching relationship between texts, this paper

proposes a Chinese semantic matching algorithm based on RoBERTa-wwm-ext [7] with Siamese interaction and fine-tuned representation (RSIFR). RoBERTa-wwm-ext is used as the baseline model for Chinese preprocessing. For the preprocessing vectors, the Siamese interaction structure with embedded soft alignment attention mechanism and BiLSTM and the LSTM-BiLSTM network structure are built, which further enhance the representation ability of vectors. Then, a model for sentence pair classification fine-tuned based on RoBERTa-wwm-ext is built and pretrained, which is used to extract fine-tuned representation vectors of sentence pairs. The MLP structure is built for the final generated vector.

The main contributions of this paper are as follows:

(1) Input the two texts independently into RoBERTa-wwm-ext model and extract the Pooler_out layer vector of the model. For this vector, a Siamese interaction structure embedded with a soft-aligned attention mechanism and BiLSTM is built to enhance the semantic interaction between the two texts.

(2) Concatenate the two texts into a single-sentence text and input it into the RoBERTa-wwm-ext model and extract the Pooler_out layer vector of the model. An LSTM-BiLSTM network layer is built for this vector, which enhances the vector's expression of textual contextual semantic information.

(3) A training model that can fine-tune the initial vector of RoBERTa-wwm-ext is constructed, and a text vector fine-tuned by label supervision is constructed, which further improves the representation of the semantic relationship between texts by the vector.

1.1. Related Work.

Semantic analysis is a fundamental task in various research fields such as text matching and ontology alignment (OA) [8]. The innovation of deep learning techniques provides a new technical support for semantic analysis tasks. Techniques such as RNN [9], CNN [10], and LSTM [11] are used to extract the features of text, which greatly improves the ability of feature vectors to characterize the semantic information of sentences. The updated iterations of the BERT series models have led to a breakthrough in the ability of vectors to express text semantics, providing research value in the field of Chinese semantic matching.

Deep neural network-based models are an important research direction in the field of Chinese semantic matching. Ranasinghe et al. used various combinations of GRU, Bi-LSTM, etc., in Siamese network structures to compare the representational power of various variants of the structure for text semantics [12]. Zhang et al. combined TF-IDF and Jaccard coefficients with CNN to improve the representation of vectors for sentence features, but lacked semantic links between different words [13]. Guo et al. analyzed the multiple semantic compositions of texts in terms of their frame structure and combined with the self-attention mechanism to enhance the representation of vectors for multiple semantic sentences, but lacked the representation of semantic relations between different texts [14]. Zhao et al.

considered two granularities of words and characters in text and built a Siamese network structure containing BiLSTM, soft alignment attention mechanism to enhance the semantic interaction between text pairs [15].

The text feature vectors extracted by the BERT model exhibit good semantic representations [6], and various variant models incorporating BERT have emerged in the research area of text semantic matching. Peinelt et al. further enhanced the vector representation by incorporating sentence topic features based on the BERT model from the perspective of analyzing information about the topic elements of a sentence [16]. Viji and Revathy combined BERT with a BiLSTM-based Siamese structure, and they fed the vectors generated by BERT into the twin network for further hierarchical training to enhance the semantic representation of the vectors [17]. Srinarasi et al. used a combination of WordNet and BERT models to represent semantic features of text, further enhancing the representation of contextual semantic information within feature vector text [9]. Cui et al. proposed a whole-word masking approach to Chinese semantic training based on the BERT family of models for the structural properties of Chinese [7]. They constructed a series of Chinese pretraining models based on BERT, ALBERT [18], RoBERTa [19], etc., and applied the models to Chinese semantic matching tasks with relatively excellent performance.

In summary, the method based on deep neural network technology is characterized by its ability to effectively extract the contextual information of text semantics, but the ability to capture the semantic interactions between different texts is insufficient. The method based on pretrained models features effective representation of internal semantic relations of texts, but it lacks semantic interactivity between different texts. In this regard, this paper combines the RoBERTa-wwm-ext Chinese pretraining model with LSTM and BiLSTM, incorporates the SA-Attention (soft alignment attention mechanism), constructs Siamese interaction structures, and combines the RoBERTa-wwm-ext fine-tuned sentence pair classification model to improve the accuracy in Chinese semantic matching tasks.

2. Methodology

2.1. Model Framework.

In this paper, we propose a Chinese semantic matching algorithm based on RSIFR. The model architecture is shown in Figure 1.

In Network Channel 1 (NC1), the two texts are independently connected to the RoBERTa-wwm-ext model to obtain the initial vector of the text. Then, a Siamese interaction structure with embedded SA-Attention_BiLSTM is built. The two initial vectors are crossed into the two Siamese channels and fused to produce the Siamese interaction type feature vector SiaVec.

In Network Channel 2 (NC2), we concatenate the two texts and feed the RoBERTa-wwm-ext model to extract the initial vector of the text. The initial vector is input to the LSTM-BiLSTM network layer to generate vector LBvec.

In Network Channel 3 (NC3), a sentence pair classification model based on RoBERTa-wwm-ext fine-tuning is
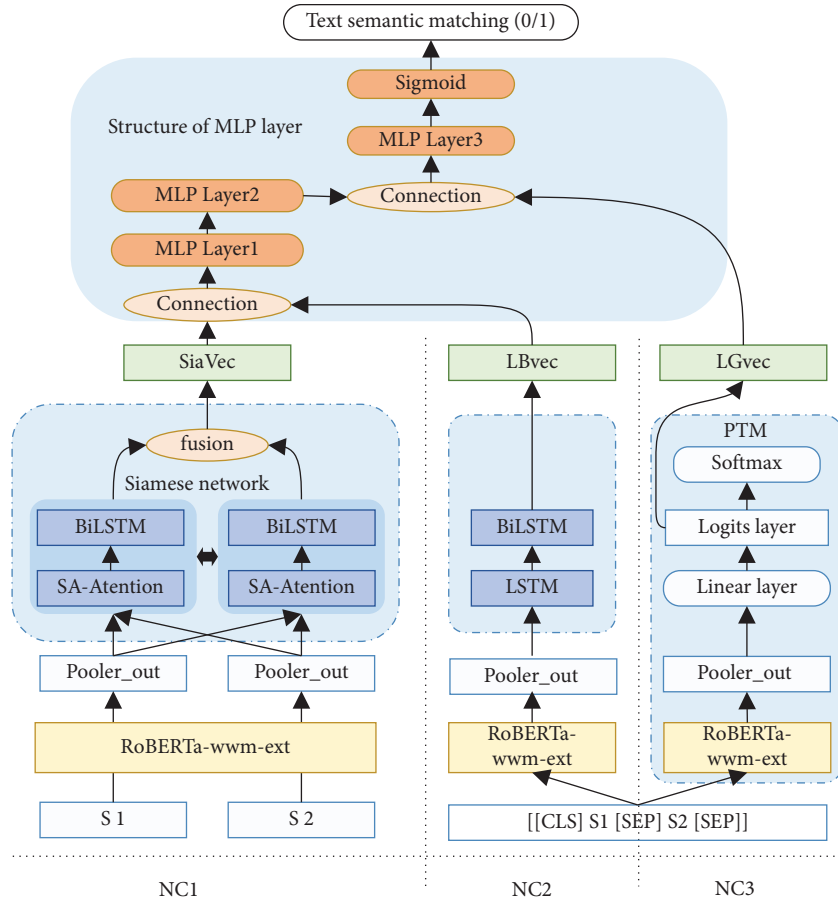
FIGURE 1: Model architecture diagram.

built, and a pretraining model PTM for sentence pair classification is generated for the dataset training, and the logit layer vector LGvec of the PTM is extracted.

At the MLP structure layer, vectors SiaVec and LBvec are connected and input to the first two fully connected layers of the MLP. Then, we concatenate the output vector with vector LGvec and feed the result into the last layer of the MLP, using sigmoid as the activation function to produce the final matching result of the sentence pair.

2.2. RoBERTa-wwm-ext Vectorization. RoBERTa-wwm-ext is a Chinese pretraining model, which adds whole-word masking (wwm) technology to the RoBERTa model and performs incremental training for large-scale Chinese data [7]. We use RoBERTa-wwm-ext as the baseline model of this model to initially extract the semantic features of Chinese texts to provide support for the downstream tasks of the model.

The two sentences S1 and S2 are input to the RoB-ERTaWE (RoBERTa-wwm-ext) model independently, and then the Pooler_out output layer vectors of the model are extracted separately. The formula is as follows:

$$S1\text{vec} = [\text{RoBERTaWE}([S1])]_{Pooler\text{-}out},$$
$$S2\text{vec} = [\text{RoBERTa}aWE([S2])]_{Pooler\text{-}out}, \tag{1}$$

where S1vec and S2vec are the initial feature vectors of text S1 and S2, respectively.

2.3. Siamese Interaction Structure. The Chinese semantic matching task is to determine whether the meanings expressed by two different Chinese sentences are consistent. Siamese network is to input two matching Chinese texts into two Siamese subchannels independently, and two independent subchannels share training weights. The Siamese structure not only achieves the training independence of the two texts but also does not ignore the information interaction between the two texts. Within the two Siamese subchannels, the BiLSTM model is used to train the contextual semantic relationships of text, and the attention mechanism is used to enhance the interaction of text semantic information. The Siamese network structure based on SA-Attention and BiLSTM not only considers the learning of similar features between two sentences but also effectively exploits the heterogeneous information between two sentences, which enhances the performance of Chinese semantic matching tasks.

The Siamese interaction structure is shown in the NC1 channel in Figure 1. In the NC1 channel, the feature vectors of the two sentences are input into two Siamese subchannels, respectively. At the same time, after the feature vector of the sentence is processed by each layer of network structure, it is

connected with the vector before processing, so as to retain the original semantic features of the text and avoid the loss of information. Finally, the vectors computed by the two subchannels are fused to produce the final sentence pair vector representation. The network processing process is shown in Figure 2.

The vectors S1vec and S2vec are crossed input to the Siamese interaction structure and first processed by SA-Attention. The attention scoring function of vectors S1vec and S2vec is as follows:

$$Score\,(S1vec, S2vec) = S1vec^T \cdot S2vec, \quad (2)$$

where $Score\,(S1vec, S2vec)$ is the attention scoring function, and then the attention distribution is calculated using the softmax function. The formula is as follows:

$$P = \frac{\exp\,(Score\,(S1vec, S2vec))}{\sum_{i=1}^{\dim} \exp\,(Score\,(S1vec, S2vec))_i}, \quad (3)$$

where P is the attention distribution function. We multiply P with the vectors S1vec and S2vec to calculate the corresponding weighted distribution. In order to avoid information loss, the initial vectors S1vec and S2vec are added to the calculation results. At the same time, the result of the addition is connected with S1vec and S2vec, respectively. The formula is as follows:

$$Avec1 = Concat\,[P \cdot S1vec + S1vec, S1vec],$$
$$Avec2 = Concat\,[P \cdot S2vec + S2vec, S2vec]. \quad (4)$$

The vectors Avec1 and Avec2 are input to the BiLSTM, and the output vectors are concatenated with the vectors generated in the previous steps, respectively. The formula is as follows:

$$SiaVec1 = Concat\,[BiLSTM\,(Avec1), Avec1, S1vec],$$
$$SiaVec2 = Concat\,[BiLSTM\,(Avec2), Avec2, S2vec]. \quad (5)$$

The vectors SiaVec1 and SiaVec2 are fused to produce the final vector representation SiaVec of the sentence pair. The formula is as follows:

$$SiaVec = [|SiaVec1 - SiaVec2|, SiaVec1 \odot SiaVec2], \quad (6)$$

where $\odot$ represents the multiplication of the corresponding terms of vectors SiaVec1 and SiaVec2. The vector SiaVec preserves the original semantic information of the text and enhances the semantic interaction between the two sentences.

2.4. LSTM-BiLSTM Network Structure.
The LSTM model captures the semantic relationships between long-distance words in text very well, and it focuses on the forward encoding relationships in text sentences. The BiLSTM model focuses on both the positive and negative directions of the text, effectively expressing the context semantic relationship of the text. Chinese text semantics has a strong positive logical relationship. Based on this, we first use LSTM model to enhance the forward logical semantic representation of

text semantics. Then, through the BiLSTM model, we pay more attention to the text semantic forward logic, but also learn the reverse semantic logical relationship of text. The LSTM-BiLSTM fusion model more effectively enhances the vector's contextual semantics for text.

The model is shown in Figure 1 for the NC2 channel. Firstly, two sentences S1 and S2 are connected to one sentence of text. The text Sen is input to the RoBERTaWE (RoBERTa-wwm-ext), and the output Pvec of the Pooler_out layer is extracted as the initial vector representation of the text. The formula is as follows:

$$Sen = [[CLS]S1[SEP]S2[SEP]],$$
$$Pvec = [RoBERTaWE\,([Sen])]_{Pooler\text{-}out.} \quad (7)$$

Then, vector Pvec is input to the LSTM layer, denoted as LSTM([Pvec]). To avoid losing information, first concatenate LSTM([Pvec]) with Pvec, and then enter the BiLSTM layer to obtain the final vector LBvec. The formula is as follows:

$$LBvec = BiLSTM\,(Concat\,[LSTM\,([Pvec]), Pvec]), \quad (8)$$

where LBvec is based on the RoBERTa-wwm-ext, which further enhances the semantic interaction within a single text and between two texts and enriches the representational information of the sentence pairs embedded in the vector.

2.5. Text Feature Representation Based on RoBERTa-wwm-ext Fine-Tuning.
The feature vectors extracted directly from the RoBERTa-wwm-ext model ignore the influence of labels on the representation of text feature vectors. A RoBERTa-wwm-ext training model with fine-tuning mechanism is constructed, and the text's eigenvector parameters are adjusted through label supervision. The feature vector extracted by this structure covers the semantic association between text pairs, which improves the Chinese semantic matching performance.

The model is shown in Figure 1 for the NC3 channel. Firstly, the output vector Pvec of the Pooler_out layer of RoBERTa-wwm-ext is input to the linear transformation layer. The formula is as follows:

$$LWvec = Pvec \cdot W^T + Bias, \quad (9)$$

where W is the weight matrix of the vector Pvec undergoing linear transformation and Bias is the bias of the function.

Then, the vector LWvec passes through the softmax activation layer, resulting in the final text pair matching result P_{LWvec}. The formula is as follows:

$$P_{LWvec} = \frac{\exp\,(LWvec)}{\sum_{i=1}^{\dim} LWvec_i}. \quad (10)$$

Supervised training is performed on the data to generate the sentence pair classification pretraining model PTM, and the logit output layer is extracted as a feature vector for the fine-tuned type of text pairs. The formula is as follows:

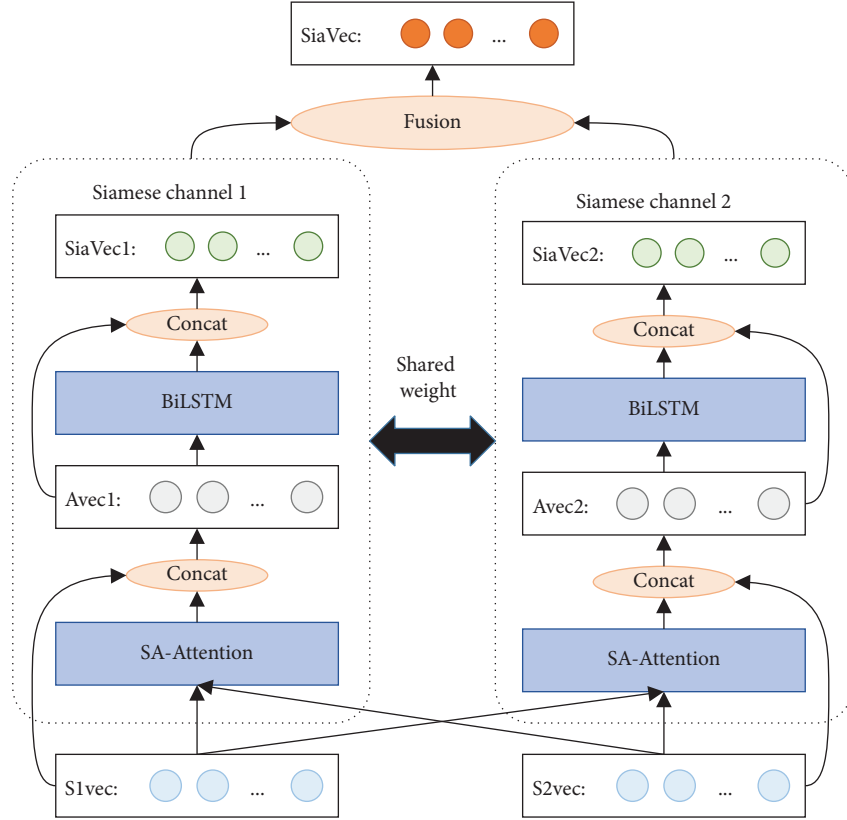$$LGvec = [PTM\,([Sen])]_{Logits}. \quad (11)$$

Figure 2: Siamese interaction structure.

The fine-tuned vector LGvec, which directly contains the semantic matching relationship between sentence pairs, plays a key role in the subsequent judgment of the matching degree of text pairs.

2.6. MLP Structure.

After the analysis in the previous sections, vector LGvec contains the semantic matching information of sentence pairs, so vector LGvec is not involved in training in the first two fully connected layers of the MLP layer to avoid the loss of matching information.

Firstly, vectors SiaVec and LBvec are connected and participate in the training of the first two fully connected layers of MLP, and R_{L2} is the output of the fully connected layer. The formula is as follows:

$$R_{L1} = \text{MLP}_{layer1}\left(Concat\left[\text{SiaVec}, \text{LBvec}\right]\right),$$
$$R_{L2} = \text{MLP}_{layer2}\left(\left[R_{L1}\right]\right). \tag{12}$$

Then, the vectors R_{L2} and LGvec are connected to participate in the training of the fully connected layer of MLP layer 3, and finally, the final matching result is output by the sigmoid activation function. The formula is as follows:

$$\begin{cases} R_{L3} = \text{MLP}_{layer3}\left(Concat\left[R_{L2}, \text{LGvec}\right]\right), \\ R = \text{MLP}_{Sigmoid}\left(\left[R_{L3}\right]\right)(R = 0: Mismatch, R = 1: Match), \end{cases} \tag{13}$$

where R is the matching result.

2.7. RSIFR Algorithm Implementation.

The algorithm implementation of the model RSIFR is mainly divided into Chinese preprocessing and Chinese semantic matching classification training. The overall algorithm framework includes the entire processing process from the initial sentence pair input to the final semantic matching result, which more clearly shows the algorithm composition of RSIFR. The algorithm process is shown in Table 1.

3. Experiment

3.1. Dataset.

The Chinese text contained in the public dataset LCQMC [20] covers a wide range of fields and is widely used in the research of Chinese semantic matching-related tasks. Therefore, in the experiments of this paper, we use LCQMC as the experimental dataset of this model. The size of the dataset is shown in Table 2.

The format of the LCQMC dataset is two Chinese texts corresponding to a 0/1 tag, with 0 indicating a semantic mismatch between the two texts and 1 indicating a match. The example of the dataset is shown in Table 3.

3.2. Ablation Experiments.

The model RSIFR in this paper is designed with three core structures NC1, NC2, and NC3, as shown in Figure 1. To prove the validity and necessity of each structure in the model, the three modules NC1, NC2, and NC3 are eliminated on the basis of RSIFR model, and the evaluation index changes of RSIFR, $RSIFR^{NC1+NC2}$, $RSIFR^{NC1+NC3}$, and $RSIFR^{NC1+NC3}$ models are

TABLE 1: RSIFR algorithm processing process.

RSIFR algorithm
1: Input: Sentence pairs S1, S2
2: Output: Semantic matching results R
3: S1, S2 independent input RoBERTa-WM, extract the pool_out layer vector SP1, SP2
4: S1, S2 connect input RoBERTa-WM, extract the pool_out layer P
5: Sentence pair classification model pretraining
6: for each epochs
7: S1, S2 input RoBERTa-WM, extract the pool_out layer vector D
8: Logits = linear (D)
9: res = softmax (logits)
10: end for
11: Training generates sentence pair classification pretraining model PTM
12: S1, S2 input PTM, extract logits layer vector LG
13: for each epochs
14: SP1, SP2 input SA-Attention, resulting in P1, P2
15: P1, P2 input BiLSTM, resulting in L1, L2
16: Concat [SP1, P1, L1] => SV1
17: Concat [SP2, P2, L2] => SV2
18: Concat [|SV1-SV2|, SV1* SV2] => SV
19: P input LSTM, resulting in PL
20: Concat [P, PL] input BiLSTM, resulting in LB
21: SV, LB, LG input MLP
22: Concat [SV, LB] Input MLP layer 1, resulting in R1. dropout (0.2)
23: R1 Input MLP layer 2, resulting in R2. dropout (0.2)
24: Concat [R2, LG] Input MLP layer 3, resulting in R3. dropout (0.2)
25: R = sigmoid ($R3$)
26: end for

experimentally compared. The evaluation index of the model utilizes the ACC and F1 values, and the results of the experiment are shown in Table 4.

In Table 4, the evaluation indexes ACC and F1 of RSIFR are greater than the evaluation indexes of each ablated model. This proves the necessity of the simultaneous existence of the three modules in the RSIFR model, and all three modules contribute to the performance improvement of the RSIFR model.

3.3. Performance Comparison Based on Different Baseline Models.
RSIFR model is a vector representation with RoBERTa-wwm-ext to initially extract the text. Cui et al. published a series of Chinese preprocessing models such as BERT-wwm at the same time [7]. We used different Chinese preprocessing models as the baseline model for Chinese preprocessing in this paper model to compare and verify the Chinese preprocessing model with the best performance. The experimental data are shown in Table 5.

The experimental results show that RoBERTa-wwm-ext is used as the baseline model for Chinese preprocessing, and the two evaluation indicators of ACC and F1 on the LCQMC dataset achieve the maximum value, and the performance is the best.

3.4. Comparison of Existing Models.
We compare the performance of RSIFR with existing models and use ACC and F1 as the evaluation indicators of the model. The data comparison is shown in Table 6.

TABLE 2: Size of LCQMC dataset.

Dataset	Total	Positive	Negative
Training	238766	138574	100192
Validation	8802	4402	4400
Test	12500	6250	6250

The existing BERT-related Chinese semantic matching model is a model derived from an improved model based on BERT. BERT-wwm is added with the method of full-word masking based on the BERT. BERT-wwm-ext has been trained incrementally in Chinese on the basis of BERT-wwm [7]. RoBERTa and MacBERT [7] are two other improved models of BERT. The two models were separately trained incrementally, resulting in Chinese semantic pretraining models such as RoBERTa-wwm-ext, RoBERTa-wwm-ext-large, MacBERT-base, and MacBERT-large [7]. They are applied to Chinese semantic matching tasks and show relatively good performance.

Among other existing models, Lattice-CNN is used to extract text semantic information from the perspective of multigranularity of text [21]; BiMPM is used to jointly capture the contextual semantic of the text from both positive and negative directions [22]; ESIM is used to use attention mechanism on text sequences to achieve inference between text sequences [23]; CATsNET is used to capture nonlocal features of text by building a Siamese network of crossattention mechanism [24]; GMN is used to build a graph structure that effectively expresses multiple textual meanings and combined with BERT [25];

Table 3: Example of LCQMC dataset.

Type	Sentence pair	Label
Positive	S1: 怎么能让玻璃更干净 (How can we make the glass cleaner) S2: 如何把玻璃擦得又干净又亮 (How to wipe the glass clean and bright)	1
Negative	S1: 你是学生还是老师 (Are you a student or a teacher) S2: 你是徐老师学生吗 (Are you Mr. Xu's student)	0

Table 4: Results of ablation experiments.

Model	ACC	$F1$
$RSIFR^{\text{NC1+NC2}}$	83.62	83.80
$RSIFR^{\text{NC1+NC3}}$	88.98	89.35
$RSIFR^{\text{NC2+NC3}}$	88.20	88.85
$RSIFR^{\text{NC1+NC2+NC3}}$ (ours)	89.56	89.67

Table 5: Experimental results based on different baseline models.

Model	ACC	$F1$
Based on BERT-base	88.90	88.84
Based on BERT-wwm	87.79	88.35
Based on BERT-wwm-ext	87.61	87.84
Based on ALBERT-base	88.50	88.20
Based on RoBERTa-wwm-ext-large	86.92	87.78
Based on RoBERTa-wwm-ext (ours)	**89.56**	**89.67**

Table 6: Comparison of existing models.

BERT-related models	ACC	$F1$	Other existing models	ACC	$F1$
BERT	86.9	—	Lattice-CNN	82.1	82.4
BERT-wwm	87.0	—	BiMPM	83.3	84.9
BERT-wwm-ext	87.1	—	ESIM	82.6	84.5
RoBERTa-wwm-ext	86.4	—	CATsNET	83.15	—
RoBERTa-wwm-ext-large	87.0	—	GMN	84.6	86.0
MacBERT-base	87.0	—	COIN	86.2	87.0
MacBERT-large	87.6	—	PERT	87.2	—
GMN-BERT	87.3	88.0	ABOEN	84.89	85.91
StyleBERT	87.9	—			
RSIFR (ours)	**89.56**	**89.67**	**RSIFR (ours)**	**89.56**	**89.67**

StyleBERT is used to combine Chinese pinyin, strokes, and other dimensions to enrich Chinese representation [26]; COIN is used for semantic alignment of different text sequences by establishing a context-aware crossattention mechanism [27]; PERT is used to establish a training method for text position substitution by combining N-gram and whole-word masking methods [28]; and ABOEN is an attention-based semantic enhancement model, which is used to extract finer-grained semantic information [29].

The two evaluation indicators of ACC and F1 of the RSIFR algorithm model achieved the maximum value in the comparison model, and the data are shown in Table 6. Experimental data show that the performance of the RSIFR algorithm model on the data LCQMC is better than the existing Chinese semantic matching model.

Before the RSIFR model worked, we locally prestored the initial vectors of the text. Subsequent text semantic analysis and text-matching tasks of the model are based on locally stored text initial vectors. In the process of model training, the time cost of initial text eigenvector initialization is saved, the training efficiency and running performance of the model are improved, and the training cost of the model is reduced.

4. Conclusions

In this paper, we propose a Chinese semantic matching algorithm based on RoBERTa-wwm-ext with Siamese interaction and fine-tuned representation. We study and design the Siamese interaction structure, the LSTM-BiLSTM network structure, and the text feature representation

structure fine-tuned based on RoBERTa-wwm-ext. The three structures generate twin interaction vectors, fully connected vectors, and fine-tuned representation vectors of sentence pairs, respectively, and a specific MLP network structure is designed for the three vector representations to obtain the final semantic matching result. The RSIFR Chinese semantic matching algorithm proposed in this paper starts from multiple dimensions of Chinese text, not only considers the contextual semantic relationship within a single text but also considers the semantic heterogeneous relationship between different texts, effectively strengthening the semantic interaction between different sentences It enhances the representation ability of vectors for textual context semantics. We show through experiments that the model proposed in this paper outperforms existing Chinese semantic matching algorithms on the public dataset LCQMC.

Data Availability

The data supporting the results of this study are public datasets. The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] K. Seki, "Cross-lingual text similarity exploiting neural machine translation models," *Journal of Information Science*, vol. 47, no. 3, pp. 404–418, 2021.

[2] C. Liu, S. Jiang, and H. Yu, "Multi-turn inference matching network for natural language inference," in *Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing*, pp. 131–143, Springer, Guilin, China, August, 2018.

[3] X. Xue and C. Jiang, "Matching sensor ontologies with multi-context similarity measure and parallel compact differential evolution algorithm," *IEEE Sensors Journal*, vol. 21, no. 21, pp. 24570–24578, 2021.

[4] X. Niu, W. Zheng, and Y. Xiao, "Short Text Similarity Computation Method Based on Feature Expansion and Siamese Network," in *Proceedings of the 2021 4th International Conference on Data Science and Information Technology*, pp. 268–272, Springer, Singapore, July, 2021.

[5] Y. Yang and C. Zhang, "Attention-based multi-level network for text matching with feature fusion," in *Proceedings of the 2021 4th International Conference on Algorithms, Computing and Artificial Intelligence*, pp. 1–7, Rochester, NY, USA, December, 2021.

[6] J. Devlin, M. W. Chang, and K. Lee, "Bert: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the NAACL-HTL*, pp. 4171–4186, Stroudsburg, PA, USA, 2019.

[7] Y. Cui, W. Che, T. Liu, B. Qin, and Z. Yang, "Pre-training with whole word masking for Chinese BERT," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3504–3514, 2021.

[8] X. Xue, "Complex Ontology Alignment for Autonomous Systems via the Compact Co-evolutionary Brain Storm Optimization Algorithm," *ISA Transactions*, vol. 99, p. 1, 2022.

[9] S. Srinarasi, R. Ram, and S. Raghavendra, "A Combination of Enhanced WordNet and BERT for Semantic Textual Similarity," in *Proceedings of the 2021 2nd International Conference on Control, Robotics and Intelligent System*, pp. 191–198, Rochester, NY, USA, October, 2021.

[10] B. Hu, Z. Lu, and H. Li, "Convolutional neural network architectures for matching natural language sentences," *Advances in Neural Information Processing Systems*, vol. 27, 2014.

[11] J. Mueller and A. Thyagarajan, "Siamese recurrent architectures for learning sentence similarity," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30, no. 1, Phoenix, AZ, USA, November, 2016.

[12] T. Ranasinghe, C. Orǎsan, and R. Mitkov, "Semantic textual similarity with Siamese neural networks," in *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019)*, pp. 1004–1011, Varna, Bulgaria, September, 2019.

[13] P. Zhang, X. Huang, Y. Wang, C. Jiang, S. He, and H. Wang, "Semantic similarity computing model based on multi model fine-grained nonlinear fusion," *IEEE Access*, vol. 9, pp. 8433–8443, 2021.

[14] S. Guo, Y. Guan, R. Li, X. Li, and H. Tan, "Frame-based multi-level semantics representation for text matching," *Knowledge-Based Systems*, vol. 232, Article ID 107454, 2021.

[15] P. Zhao, W. Lu, and Y. Li, "Chinese semantic matching with multi-granularity alignment and feature fusion," in *Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Shenzhen, China, July, 2021.

[16] N. Peinelt, D. Nguyen, and M. Liakata, "tBERT: topic models and BERT joining forces for semantic similarity detection," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 7047–7055, Red Hook, NY, USA, January, 2020.

[17] D. Viji and S. Revathy, "A hybrid approach of weighted fine-tuned bert extraction with deep siamese bi-lstm model for semantic text similarity identification," *Multimedia Tools and Applications*, vol. 81, pp. 1–27, 2022.

[18] Z. Lan, M. Chen, and S. Goodman, "Albert: a lite bert for self-supervised learning of language representations," in *Proceedings of the ICLR*, Ababa, Ethiopia, April 2020.

[19] Y. Liu, M. Ott, and N. Goyal, "Roberta: a robustly optimized bert pretraining approach," 2019, https://arxiv.org/abs/1907.11692.

[20] X. Liu, Q. Chen, and C. Deng, "Lcqmc: a large-scale Chinese question matching corpus," in *Proceedings of the 27th International Conference on Computational Linguistics*, pp. 1952–1962, Stroudsburg, PA, USA, August 2018.

[21] Y. Lai, Y. Feng, and X. Yu, "Lattice Cnns for matching based chinese question answering," in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 6634–6641, Phoenix, AZ, USA, July, 2019.

[22] Z. Wang, W. Hamza, and R. Florian, "Bilateral multi-perspective matching for natural language sentences," in *Proceedings of the 26th International Joint Conference on Artificial Intelligence*, pp. 4144–4150, Macao, China, February, 2019.

[23] Q. Chen, X. Zhu, and Z. Ling, "Enhanced LSTM for natural language inference," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, vol. 1, pp. 1657–1668, Stroudsburg, PA, USA, July 2017.

[24] Z. Wang, X. Zhang, and Y. Tan, "Chinese sentences similarity via cross-attention based siamese network," 2021, https://arxiv.org/abs/2104.08787.

[25] L. Chen, Y. Zhao, and B. Lyu, "Neural graph matching networks for Chinese short text matching," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 6152–6158, Rochester, NY, USA, January, 2020.

[26] C. Lv, H. Zhang, and X. Du, "StyleBERT: chinese pre-training by font style information," 2022, https://arxiv.org/abs/2202.09955.

[27] Z. Hu, Z. Fu, and Y. Yin, "Context-aware interaction network for question matching," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp. 3846–3853, Red Hook, NY, USA, April, 2021.

[28] Y. Cui, Z. Yang, and T. Liu, "PERT: pre-training bert with permuted language model," 2022, https://arxiv.org/abs/2203.06906.

[29] H. Zhang, H. Zhang, and X. Lu, "Attention-based overall enhance network for Chinese semantic textual similarity measure," *Journal of Applied Science and Engineering*, vol. 25, no. 2, pp. 287–295, 2021.