Chinese Sentence Semantic Matching Based on Multi-Granularity Fusion Model

Xu Zhang[1], Wenpeng Lu[1(✉)], Guoqiang Zhang[2], Fangfang Li[3], and Shoujin Wang[4]

[1] School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China
Xuzhang.p@foxmail.com, Wenpeng.Lu@qlu.edu.cn
[2] Centre for Audio, Acoustics and Vibration, University of Technology Sydney, Sydney, Australia
Guoqiang.Zhang@uts.edu.au
[3] oOh! Media, Sydney, Australia
Fangfang.Li@oohmedia.com.au
[4] Department of Computing, Macquarie University, Sydney, Australia
Shoujin.Wang@mq.edu.au

Abstract. Sentence semantic matching is the cornerstone of many natural language processing tasks, including Chinese language processing. It is well known that Chinese sentences with different polysemous words or word order may have totally different semantic meanings. Thus, to represent and match the sentence semantic meaning accurately, one challenge that must be solved is how to capture the semantic features from the multi-granularity perspective, e.g., characters and words. To address the above challenge, we propose a novel sentence semantic matching model which is based on the fusion of semantic features from character-granularity and word-granularity, respectively. Particularly, the multi-granularity fusion intends to extract more semantic features to better optimize the downstream sentence semantic matching. In addition, we propose the equilibrium cross-entropy, a novel loss function, by setting mean square error (MSE) as an equilibrium factor of cross-entropy. The experimental results conducted on Chinese open data set demonstrate that our proposed model combined with binary equilibrium cross-entropy loss function is superior to the existing state-of-the-art sentence semantic matching models.

Keywords: Sentence semantic matching · Multi-granularity fusion · Equilibrium cross-entropy

1 Introduction

Sentence semantic matching plays a key role in many natural language processing tasks such as question answering (QA), natural language inference (NLI), machine translation (MT), etc. The key of sentence semantic matching is to

calculate the semantic similarity between given sentences from multiple text segmentation granularity such as character, word and phrase. Currently, the commonly used text segmentation is in word granularity only, especially for Chinese. However, many researchers have realized that a text can be viewed from not only word granularity but also the others.

In word granularity, many deep learning based sentence semantic matching models have been proposed, such as DeepMatch$_{tree}$ [18], ARC-II [5], Match-Pyramid [12], Match-SRNN [16], etc. However, these word-granularity models are unable to fully capture the semantic features embedded in sentences, sometimes even produce noise and thus hurt the performance of sentence matching. Eventually, more and more researchers turn to design semantic matching strategy combing word and phrase granularity, such as MultiGranCNN [24], MV-LSTM [15], MPCM [22], BiMPM [21], DIIN [3]. These models somehow overcome the word-granularity modelling limitations, however, they still cannot thoroughly solve the issue of semantic loss in the process of sentence encoding, especially for Chinese corpus which are usually with rich semantic features.

Similarly for Chinese sentence semantic matching task, many researchers attempt to mix words and characters together into a simple sequence. For example, multi-granularity Chinese word embedding [23] and lattice CNNs for QA [7] have achieved great performance. However, most Chinese characters cannot be treated as independent words or phrases as these works did. This is because the simple combining of characters or words together, or encoding characters according to character lattice may easily lose the meaning that is embedded in the corresponding character.

In order to capture the sentence features from both character and word perspectives more deeply and comprehensively, we propose a new sentence semantic matching model with multi-granularity fusion. The semantic features of the text are obtained from the character and word perspectives respectively, and the more critical semantic information in the text is captured through the superposition effect of the two features. Our model significantly improves the representation of textual features. Moreover, for most existing deep learning applications, cross-entropy is a commonly used loss function to train the models. We design a novel loss function, which utilizes mean square error (MSE) as an equilibrium parameter to strengthen and enhance cross-entropy with the ability to distinguish the fuzzy classification boundary, which greatly improves the performance of our model.

Our contributions are summarized as follows:

– We propose a novel sentence encoding method named multi-granularity fusion model to better capture semantic features via the integration of multi-granularity encoding.
– We propose a novel deep neural architecture for sentence semantic matching task, which includes embedding layer, multi-granularity fusion encoding layer, matching layer and prediction layer.

– We propose a new loss function integrating equilibrium parameter into cross-entropy function. MSE is introduced as the equilibrium parameter to construct the binary equilibrium cross-entropy loss.
– Our source code is publicly available[1]. Our work may provide a reference for researchers in NLP community.

The rest of the paper is structured as follows. We introduce the related work about sentence semantic matching in Sect. 2, and propose multi-granularity fusion model in Sect. 3. Section 4 demonstrates the empirical experimental results, followed by the conclusion in Sect. 5.

2   Related Work

Semantic matching in short text is the basis of natural language understanding tasks. Its improvement will help advance the progress of natural language understanding tasks. A lot of work has put great efforts into the semantic matching in short texts [3,10,16,20,21,25].

With the continuous development of deep learning, it is difficult to further obtain the text semantic information only depending on designing the models with more complex and deep architecture. The researchers then begin to consider obtaining more semantic features from texts on different granularity. In the matching process, both the sentence and the word, phrase perspectives are considered. The results of multi-faceted feature matching are combined to get better results [1,15,19,21,23,24]. Yin et al. propose MultiGranCNN to first obtain text features on different granularity such as words, phrases, and sentences, and then concatenate these text features and calculate the similarity between the two sentences [24]. Wan et al. propose MV-LSTM method similar to MultiGranCNN, which can capture long-distance and short-distance dependencies simultaneously [15]. MIX is a multi-channel convolutional neural network model for text matching, with additional attention mechanisms on sentences and semantic features [1]. MIX compares text fragments on varied granularity to form a series of multi-channel similarity matrices, which are then crossed with another set of carefully designed attention matrices to expose the rich structure of sentences to a deep neural network. Though all the above methods perform feature representation for the same text on word, phrase and sentence granularity simultaneously, they still ignore the influence of features on other granularity, such as character. In order to solve this problem in Chinese language, we generate corresponding text vectors, extracting the character-granularity and the corresponding word-granularity features separately. The feature on each granularity is captured from the corresponding text sequence.

Most tasks in natural language processing field can be considered as classification problems. For classification tasks, the most commonly used loss function in deep learning methods is cross-entropy. In view of the related tasks in computer vision, a series of loss functions based on optimization have been proposed

---

[1] https://github.com/XuZhangp/MGF.

to improve face recognition [2,8,17], image segmentation [11,13,14] and other tasks. Compared with computer vision, there is few related work on reconstructing loss function for a specific task in natural language processing field. Kriz et al. present a customized loss function to replace the standard cross-entropy during training, which takes the complexity of content words into account [6]. They propose a metric that modifies cross-entropy loss to up weight simple words and down weight more complex words for sentence simplification. Besides, Hsu et al. introduce the inconsistency loss function to replace cross-entropy loss in text extraction and summarization [4]. To better distinguish the classification results, Zhang et al. modify the cross-entropy loss function and apply it on the text matching task [25]. Inspired by the work, we propose a new loss function, where MSE is used as the balance factor to enhance the cross-entropy loss function. It can strengthen the ability to distinguish the fuzzy classification boundary in the training process and improve classification accuracy.

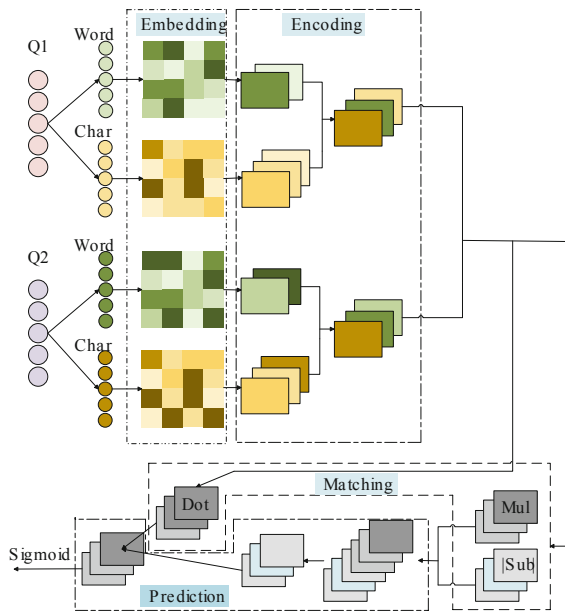## 3   Multi-Granularity Fusion Model

3.1   Model Architecture

Fig. 1. Model architecture of sentence matching

As shown in Fig. 1, our proposed model architecture includes a multi-granularity embedding layer, a multi-granularity fusion encoding layer, a matching layer and a prediction layer. First, we embed the input sentences from both character and word perspectives through the multi-granularity embedding layer. Then,

the output of multi-granularity embedding layer is transmitted to the multi-granularity fusion encoding layer to extract two streams of semantic features on the character and word granularity, respectively. When the semantic feature extraction is complete, the semantic feature is fed to the matching layer to generate a final matching representation of the input sentences, which is further transferred to a Sigmoid function to judge their matching degree in the prediction layer.

3.2 Multi-Granularity Embedding Layer

For Chinese text, after sentence segmentation from character and word perspectives, we obtain two sentence sequences based on character granularity and word granularity. By the multi-granularity embedding layer, the original sentence sequences are converted to the corresponding vector representations, respectively. In this embedding layer, we utilize the pre-trained embeddings, which are trained with Word2Vec on the target data set.

3.3 Multi-Granularity Fusion Encoding Layer

In this subsection, we introduce our key contribution module which named multi-granularity fusion encoding layer to improve the semantic encoding performance. This model integrates and considers the word vector and character vector comprehensively, which are depended on its own text sequence respectively.
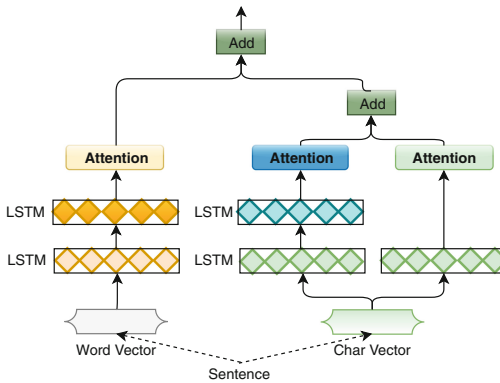
Fig. 2. Multi-Granularity Fusion Encoding

As shown in Fig. 2, for the input sentence, we use different encoding methods to generate the character-granularity sentence vectors and the word-granularity sentence vectors. Aiming at the word-granularity sentence vector, we use two LSTMs for sequential encoding, then introduce the attention mechanism on deep

feature extraction. Meanwhile, aiming at the character-granularity sentence vector, we use the same encoding method, which is similar with the word-granularity sentence vector. Moreover, for the character-granularity sentence vectors, we supplement a single layer of LSTM for encoding and then use the attention mechanism for deep feature extraction. For the above two encoding results on character granularity, we add them together to obtain more accurate semantic representation information on the character granularity.

As shown in Fig. 2, by the above operations on character-granularity and word-granularity sentence vectors, we can obtain semantic feature information on two perspectives. In order to capture more semantic features and understand the sentence semantic meaning more deeply, we add the sentence vectors from two perspectives together.

With this multi-granularity fusion encoding layer, the complex semantic features of the sentences are captured from the character and word perspectives respectively, and the more critical and important semantic information in the sentences are obtained through the superposition effect of the two features. This model can significantly improves the representation of sentence features.
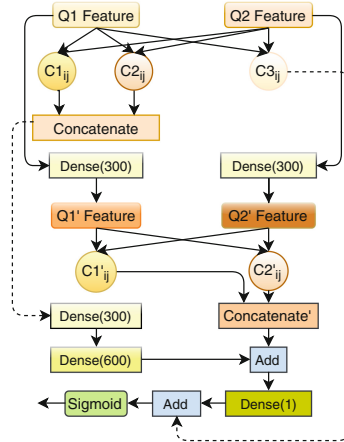
3.4 Interaction Matching Layer

Fig. 3. Interaction Matching

The multi-granularity fusion encoding layer outputs the semantic feature vectors (Q1 Feature and Q2 Feature) for the sentences Q1 and Q2, which are transferred to interaction matching layer, as shown in Fig. 3.

In the interaction matching layer, we utilize multiple calculation methods to hierarchically compare the similarity of the semantic feature vectors for sentences Q1 and Q2. The initial operations are described as follows:

$$\overrightarrow{C1}_{ij} = |\overrightarrow{Q1}_{ij} - \overrightarrow{Q2}_{ij}| \tag{1}$$

$$\overrightarrow{C2}_{ij} = \overrightarrow{Q1}_{ij} \times \overrightarrow{Q2}_{ij} \tag{2}$$

$$\overrightarrow{C3}_{ij} = \overrightarrow{Q1}_{ij} \cdot \overrightarrow{Q2}_{ij} \tag{3}$$

$$\overrightarrow{Concatenate} = [\overrightarrow{Q1}_{ij}, \overrightarrow{Q2}_{ij}] \tag{4}$$

As shown in Fig. 3, the sentence features are hierarchically matched. The input Q1 and Q2 features are handled by a full connected dense layer to generate the Q1′ and Q2′ features, which are processed and matched further with Eq. (5) and Eq. (6), whose outputs are concatenated together with Eq. (7).

$$\overrightarrow{C1'}_{ij} = |\overrightarrow{Q1'}_{ij} - \overrightarrow{Q2'}_{ij}| \tag{5}$$

$$\overrightarrow{C2'}_{ij} = \overrightarrow{Q1'}_{ij} \times \overrightarrow{Q2'}_{ij} \tag{6}$$

$$\overrightarrow{Concatenate'} = [\overrightarrow{Q1'}_{ij}, \overrightarrow{Q2'}_{ij}] \tag{7}$$

The feature representation $\overrightarrow{Concatenate}$ obtained with Eq. (4) is further extracted using two dense layers, whose dimensions are 300 and 600 respectively. Then, we add this transformed representation and another feature representation $\overrightarrow{Concatenate'}$ obtained with Eq. (7) together to generate a combined representation, followed by a dense layer whose dimension is 1. Finally, the output of the last dense layer is added to $\overrightarrow{C3}_{ij}$ obtained with Eq. (3) to generate the final matching representation of input sentences, which is further sent to the Sigmoid function to judge their matching degree in the prediction layer.

3.5   Equilibrium Cross-Entropy Loss Function

In most classification tasks, the cross-entropy loss function shown in Eq. (8), is usually the first choice. In our work, aiming to solve the difficulty of cross-entropy loss function on the fuzzy classification boundary, we try to make some modifications on cross-entropy so as to make the classification more effectively. we propose equilibrium cross-entropy by setting MSE as an equilibrium factor of cross-entropy. It can improve the accuracy when the classification boundary is fuzzy.

$$L_{crossentropy} = -\sum_{i=1}^{n}(y_{true} \log y_{pred} + (1 - y_{true}) \log(1 - y_{pred})) \tag{8}$$

As shown in Eq. (9), We use MSE as the equilibrium factor.

$$L_{mse} = \frac{1}{2n}\sum_{i=1}^{n}(y_{true} - y_{pred})^2 \tag{9}$$

By using MSE as equilibrium factor in the equilibrium loss function shown in Eq. (10), the loss function can strengthen its ability to distinguish the fuzzy boundary and eliminate the blurring phenomenon in classification tasks.

$$Loss = -\sum_{i=1}^{n}(L_{mse} * y_{true} \log y_{pred} + (1 - L_{mse}) * (1 - y_{true}) \log(1 - y_{pred})) \tag{10}$$

4   Experiments and Results

4.1   Dataset

Our methods are compared with the-state-of-art methods on the public dataset, i.e., LCQMC. It's a large-scale Chinese question matching corpus released by Liu et al. [9], which focuses on intent matching rather than paragraph matching. We use the same proportion ratio to split the dataset into training, validation and test parts, as mentioned in [9,25]. We choose a set of examples from LCQMC to introduce the text semantic matching task, shown in Table 1. From the examples, we can learn that if two sentences are matched, they should be similar in intention.

Table 1. Examples in LCQMC Corpus.

Sentence Pairs	Semantic Match?
S1:哲学的物质范畴和自然科学的物质范畴是什么关系	
EN: What is the relation between the physical category of	
philosophy and that of natural science?	
S2:简述哲学物质范畴与自然科学的物质概念关系	
EN: Describe the relation between physical category of	
philosophy and the meaning cf material in natural science	YES
S3:我对这些并不感冒是什么意思	
EN: If someone says that I'm not interested at these things,	
what does he means?	
S4:我对你并不感冒是什么意思?	
EN: If someone says that I'm not interested at you,	
what does he means?	NO

4.2   Experimental Setting

We implement our multi-granularity fusion model architecture for sentence semantic matching with Python based on Keras and Tensorflow framework. All the experiments are performed in a ThinkStation P910 Workstation with 192GB memory and one 2080Ti GPU. After testing a variety number of multi-granularity embedding layer, we empirically set its dimensionality to 300. The number of units in multi-granularity fusion encoding layer is set to 300. In the Interaction matching layer, the widths of the dense layers are shown in Fig. 3. In addition, the last dense layer utilizes sigmoid as the activation function and the other dense layers use relu. And in the multi-granularity fusion layer, we set dropout rate to 0.5. In the optimization, the epochs number is 200 and batch size is 512. We set up the early stopping mechanism. After 10 epochs, if the accuracy is not improved on the validation set, the training process will automatically stop and verify the model's performance on the test set.

4.3   Baseline Methods

On LCQMC dataset, Liu et al. [9] and Zhang et al. [25] have realized nine relevant and representative state-of-the-art methods, which are used as the baselines to evaluate our model.

– **Unsupervised Methods:** Some unsupervised matching methods based on word mover distance (WMD), word overlap (C_{wo}), n-gram overlap (C_{ngram}), edit distance (D_{edt}) and cosine similarity respectively (S_{cos}) [9].
– **Supervised Methods:** Some unsupervised matching methods based on convolutional neural network (CNN), bi-directional long short term memory (BiLSTM), bilateral multi-Perspective matching (BiMPM) [9,21] and deep feature fusion model (DFF) [25].

4.4   Performance Evaluation

A comparison of our work with the baseline methods, is shown in Table. 2, where the first fourteen rows are from (Liu et al., 2018) [9] and next two rows are from (Zhang et al., 2019) [25]. The most important indicators for sentence semantic matching task are F_1-score and accuracy. As in Table. 2, MGF surpasses the-state-of-art models on LCQMC significantly, which demonstrates the superiority of MGF.

Compared with the unsupervised methods, i.e., WMD_{char}, WMD_{word}, C_{wo}, C_{ngram}, D_{edt}, S_{cos}, our model MGF improves the precision metric by 14.39%, 16.99%, 20.29%, 29.09%, 34.89%, 21.29%, recall by 11.7%, 14.3%, 9.3%, 3.6%, 6.5%, 4.2%, F_1-score by 13.32%, 15.92%, 16.12%, 20.72%, 26.22%, 15.12% and accuracy by 15.23%, 25.83%, 15.13%, 24.63%, 33.53%, 15.53%. We can see that the improvement of our proposed model is very prominent. Compared with the unsupervised method, the proposed MGF model is a supervised one, which can use the error between the real label and the prediction to carry out backpropagation to correct and optimize the massive parameters in neural network. Besides, MGF can obtain more feature expressions through deep feature encoding. These properties gives MGF the abilities to surpass the unsupervised methods greatly.

Compared with the basic neural network methods, i.e., $CBOW_{char}$, $CBOW_{word}$, CNN_{char}, CNN_{word}, $BiLSTM_{char}$, $BiLSTM_{word}$, our model MGF improves the precision metric by 14.89%, 13.49%, 14.29%, 12.99%, 13.99%, 10.7%, recall by 10.1%, 3%, 7.3%, 8.3%, 1.9%, 3.6%, F_1-score by 12.92%, 9.32%, 11.52%, 11.02%, 9.22%, 7.8%, and accuracy by 15.23%, 12.13%, 14.03%, 13.03%, 12.33%, 9.73%. Though MGF is constructed by these basic neural network methods, it is equipped with a deeper network structure. Therefore, richer and deeper semantic features can be extracted to make the performance of our model more prominent.

Compared with the advanced neural network methods, i.e., $BiMPM_{char}$, $BiMPM_{word}$, DFF_{char}, DFF_{word}, our model MGF improves the precision metric by 3.79%, 3.69%, 2.81%, 3.7%, recall by -1%, -0.6%, -0.98%, -1.18%, F_1-score by 1.72%, 1.82%, 1.21%, 1.66% and accuracy by 2.43%, 2.53%, 1.68%,

Table 2. Experiments on LCQMC. *char* means embeddings are character-based and *word* means word-based.

Methods	Precision	Recall	F_1-score	Accuracy
WMD_{char}	67.0	81.2	73.4	70.6
WMD_{word}	64.4	78.6	70.8	60.0
C_{wo}	61.1	83.6	70.6	70.7
C_{ngram}	52.3	89.3	66.0	61.2
D_{edt}	46.5	86.4	60.5	52.3
S_{cos}	60.1	88.7	71.6	70.3
$CBOW_{char}$	66.5	82.8	73.8	70.6
$CBOW_{word}$	67.9	89.9	77.4	73.7
CNN_{char}	67.1	85.6	75.2	71.8
CNN_{word}	68.4	84.6	75.7	72.8
$BiLSTM_{char}$	67.4	91.0	77.5	73.5
$BiLSTM_{word}$	70.6	89.3	78.92	76.1
$BiMPM_{char}$	77.6	93.9	85.0	83.4
$BiMPM_{word}$	77.7	93.5	84.9	83.3
DFF^o_{char}	78.58	93.88	<u>85.51</u>	<u>84.15</u>
DFF^o_{word}	77.69	94.08	85.06	83.53
Our Models				
MGF	81.39	92.90	**86.72**	**85.83**

2.3%. BiMPM is a bilateral multi-perspective matching model, which utilizes BiLSTM to learn the sentence representation and implements four strategies to match the sentences from different perspectives [21]. DFF is a deep feature fusion model for sentence representation, which is integrated into the popular deep architecture for SSM task [25]. Compared with BiMPM and DFF, MGF realizes multi-granularity fusion encoding, which considers both character and word perspectives for the whole text. MGF can capture more comprehensive and complicated features, which leads to a better performance than the others.

5   Conclusions

To better address the Chinese sentence matching problem better, we put forward a new sentence matching model, i.e., multi-granularity fusion model, which takes both Chinese word-granularity and character-granularity into account. Specifically, we integrate word and character embedding representations together, and capture more hierarchical matching features between sentences. In addition, to solve the fuzzy boundary problem in the classification process, we use MSE as an equilibrium factor to improve the cross-entropy loss function. Extensive experiments on the real-world data set, i.e., LCQMC, have clearly shown that

our model outperforms the existing state-of-the-art methods. In future, we will introduce more features on different granularity. i.e., n-grams and phrases, etc., to encode and represent the sentences more comprehensively, and try to further improve semantic matching performance.

References

1. Chen, H., et al.: Mix: multi-channel information crossing for text matching. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 110–119. ACM (2018)
2. Deng, J., Guo, J., Xue, N., Zafeiriou, S.: Arcface: additive angular margin loss for deep face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4690–4699 (2019)
3. Gong, Y., Luo, H., Zhang, J.: Natural language inference over interaction space. arXiv preprint arXiv:1709.04348 (2017)
4. Hsu, W.T., Lin, C.K., Lee, M.Y., Min, K., Tang, J., Sun, M.: A unified model for extractive and abstractive summarization using inconsistency loss. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, pp. 132–141 (2018)
5. Hu, B., Lu, Z., Li, H., Chen, Q.: Convolutional neural network architectures for matching natural language sentences. In: Proceedings of Advances in Neural Information Processing Systems, pp. 2042–2050 (2014)
6. Kriz, R., et al.: Complexity-weighted loss and diverse reranking for sentence simplification. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 3137–3147 (2019)
7. Lai, Y., Feng, Y., Yu, X., Wang, Z., Xu, K., Zhao, D.: Lattice CNNS for matching based Chinese question answering **33**, pp. 6634–6641 (2019)
8. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988 (2017)
9. Liu, X., et al.: LCQMC: a large-scale Chinese question matching corpus. In: Proceedings of the 27th International Conference on Computational Linguistics, pp. 1952–1962 (2018)
10. Mueller, J., Thyagarajan, A.: Siamese recurrent architectures for learning sentence similarity. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pp. 2786–2792 (2016)
11. Obukhov, A., Georgoulis, S., Dai, D., Van Gool, L.: Gated CRF loss for weakly supervised semantic image segmentation. arXiv preprint arXiv:1906.04651 (2019)
12. Pang, L., Lan, Y., Guo, J., Xu, J., Wan, S., Cheng, X.: Text matching as image recognition. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pp. 2793–2799 (2016)
13. Tang, M., Djelouah, A., Perazzi, F., Boykov, Y., Schroers, C.: Normalized cut loss for weakly-supervised CNN segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1818–1827 (2018)

14. Tang, M., Perazzi, F., Djelouah, A., Ben Ayed, I., Schroers, C., Boykov, Y.: On regularized losses for weakly-supervised CNN segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 507–522 (2018)
15. Wan, S., Lan, Y., Guo, J., Xu, J., Pang, L., Cheng, X.: A deep architecture for semantic matching with multiple positional sentence representations. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, pp. 2835–2841 (2016)
16. Wan, S., Lan, Y., Xu, J., Guo, J., Pang, L., Cheng, X.: Match-SRNN: modeling the recursive matching structure with spatial RNN. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, pp. 2922–2928 (2016)
17. Wang, F., Cheng, J., Liu, W., Liu, H.: Additive margin softmax for face verification. IEEE Signal Process. Lett. **25**(7), 926–930 (2018)
18. Wang, M., Lu, Z., Li, H., Liu, Q.: Syntax-based deep matching of short texts. In: Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, pp. 1354–1361
19. Wang, S., Cao, L.: Inferring implicit rules by learning explicit and hidden item dependency. IEEE Trans. Syst. Man Cybern. Syst. **1**, 1–12 (2017)
20. Wang, S., Hu, L., Wang, Y., Cao, L., Sheng, Q.Z., Orgun, M.: Sequential recommender systems: challenges, progress and prospects. In: Proceedings of the 28th International Joint Conference on Artificial Intelligence, pp. 6332–6338 (2019)
21. Wang, Z., Hamza, W., Florian, R.: Bilateral multi-perspective matching for natural language sentences. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 4144–4150 (2017)
22. Wang, Z., Mi, H., Hamza, W., Florian, R.: Multi-perspective context matching for machine comprehension. arXiv preprint arXiv:1612.04211 (2016)
23. Yin, R., Wang, Q., Li, P., Li, R., Wang, B.: Multi-granularity Chinese word embedding. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp. 981–986 (2016)
24. Yin, W., Schütze, H.: MultiGranCNN: an architecture for general matching of text chunks on multiple levels of granularity. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, vol. 1, pp. 63–73 (2015)
25. Zhang, X., Lu, W., Li, F., Peng, X., Zhang, R.: Deep feature fusion model for sentence semantic matching. Comput. Mater. Continua. **61**, 601–616 (2019)