Integrating Word Knowledge and WordNet for Neural Metaphor Detection

Z.M. Tian[1,*], Q.Z. Wang[1], Z.H. Huang[1]

Supervisor: Prof. W.F. Su

[1] *Computer Science and Technology, Division of Science and Technology, BNU-HKBU United International College*

** Corresponding student author. Tel: +86-13172650285, E-mail: m730026102@mail.uic.edu.cn*

Abstract

- Metaphor expressions appear pervasively in daily life as well as in many literary works. A large number of implicit metaphorical expressions in any language has become a problematic task in Natural Language Processing.
- To the best of our knowledge, the state-of-the-art metaphor detection methods exploit various neural network methods, whereas most current novel neural network metaphor detection models **ignore the significant impact of knowledge information among words**. Observing that one model is likely to have a higher performance to detect metaphors if the model has more knowledge integration, our work tries to include knowledge information into the metaphor detection procedure.
- Our contribution includes augmenting knowledge information of words, incorporating knowledge base and transferring task into question answering task, which is capable of improving the performance of metaphor detection greatly.
- **An effective end-to-end metaphor detection framework that concatenates vectors from the transformer model with knowledge integration and graph embedding is proposed**. Comparing to the top-level performance of metaphor detection models on VUA and MOH-X datasets, our experimental results prove that our model achieves a competitive outcome.

Task Description

- The metaphor detection task can be described as a sequence labelling task. Given a specific sentence as input, the results should give the label of metaphorical or literal for each word as output.

Alice went to a *zoo*.
The classroom was a *zoo*.

Fig 1: metaphor detection task example. The first sentence contains all literal meaning. For the second sentence, the model output should be [0 0 0 1], which represents the word "zoo" is a metaphorical word, whereas the rest of the words are literal meaning words.

Inspiration and Intuition

- For a normal human being, one understands the metaphorical word in a sentence through two approaches: 1) **Existed knowledge information**: if one person has already known what a specific word means, he may understand this word from brain perception. 2) **Looking up in a dictionary**: if one does not know in the brain, one may look up in dictionary for further understanding.
- We develop our model inspired from the mentioned two ways. The former way is constructed through enhanced word embedding with knowledge augmentation, and the latter way is resolved by WordNet Graph Embedding.

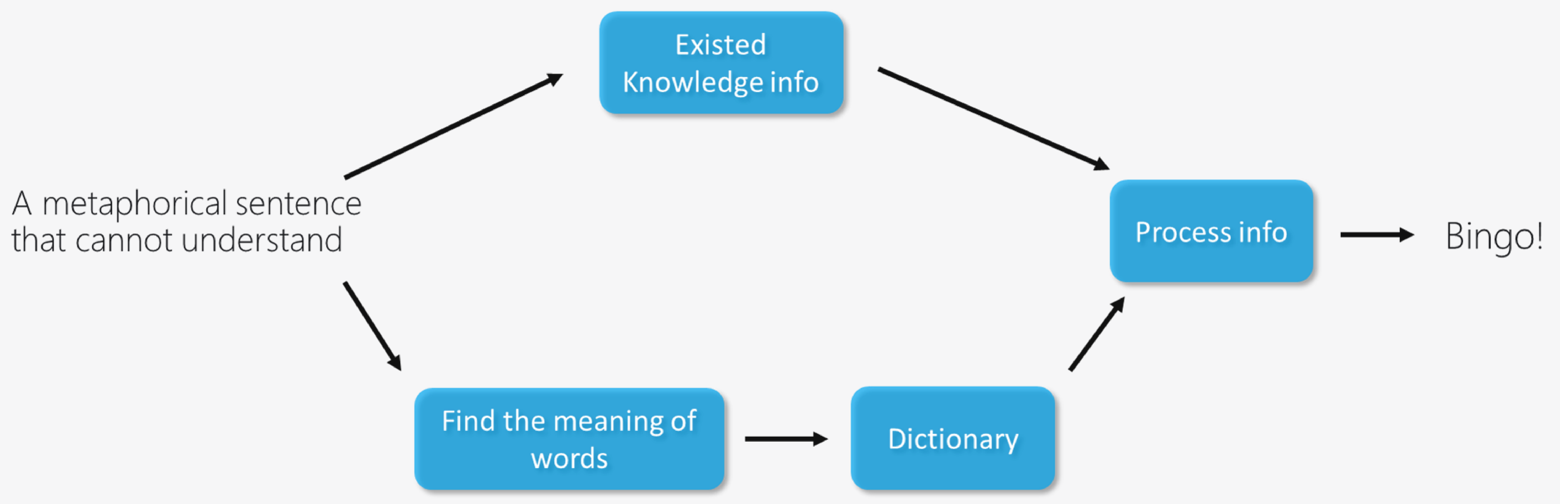

Fig 2: the procedure for our model to understand metaphor

Knowledge Integration Neural Metaphor Detection Model

The model combines context, query word, POS, TAG and knowledge as token embedding. Use token embedding, segment embedding and position embedding as the input of two transformer models: ERNIE for capturing global context and another ERNIE for local context. The output of transformer models are processed by average pooling layer and concatenate with BiLSTM on WordNet Graph Embedding. The concatenation result processes through two layers of MLP and use Softmax to calculate the probability.

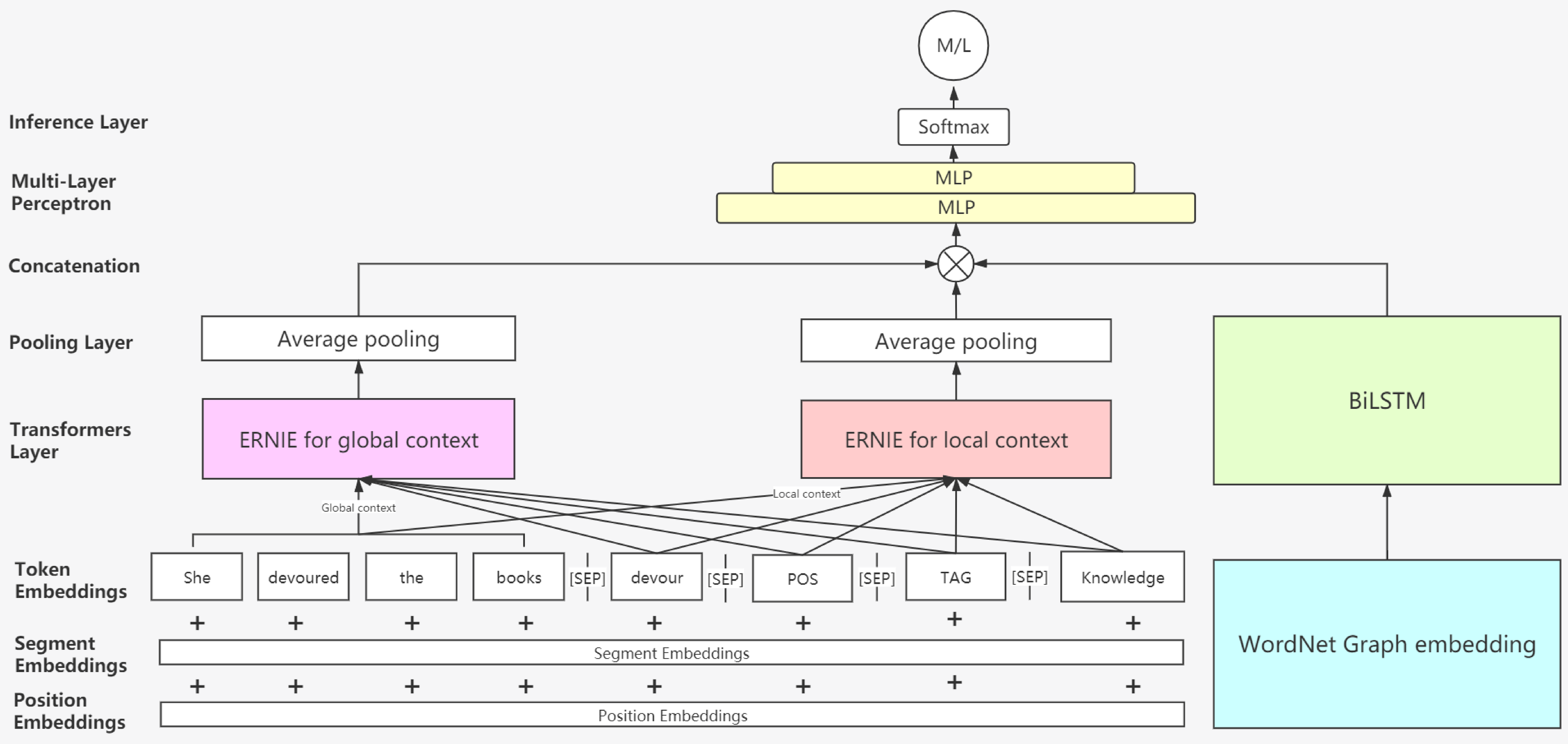

Fig 3: Model overview. The model concatenates results from two ERNIE modules and BiLSTM.

Enhanced Word Embedding

- Inspired by McCann et al. [1], we transfer the sequence labelling task into a question-answering task. Each word's additional information is provided, including POS (Part Of Speech), TAG (Detail information of POS), and knowledge. With this kind of augmentation, given the word extra information to answer a specific question may perform better than solely provide the context information. For instance, the question-answering task's procedure becomes:
- (1) Given the context, "She **devoured** the book."
- (2) Enriched with the information of the word "devoured" with knowledge "destroy completely; enjoy avidly, eat up completely", and "devoured" is a verb with past tense.
- (3) Ask the word "devoured" with the above information if it is "metaphorical" or "literal".

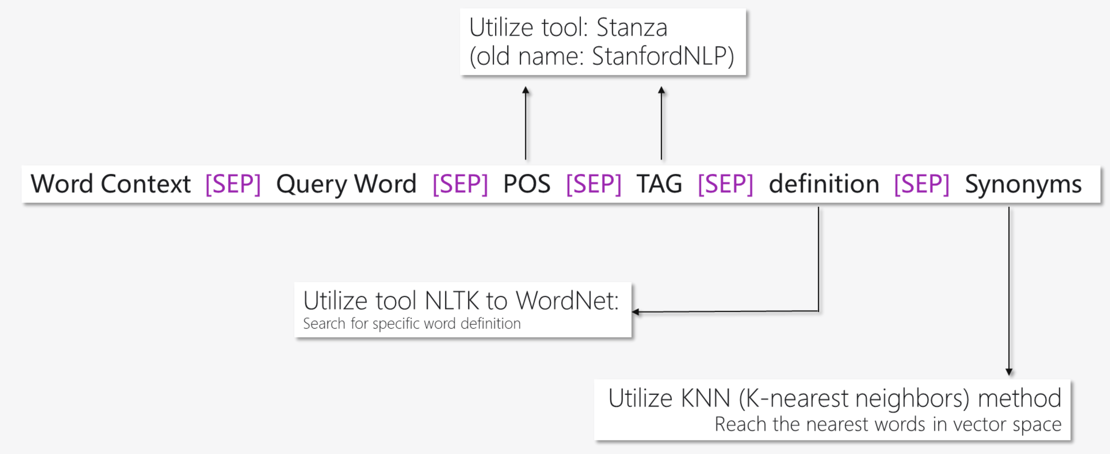

Fig 3: Enhanced Word Embedding for model understanding

Knowledge Integration Strategy

- After enhanced embedding, the embedding should transfer to fit the transformer. The transformer word embedding includes token embedding, segment embedding, and position embedding. Those embedding are fed into a double parallel transformer encoder model — processing both global and local context for the metaphor detection task.
- The transformer our model used is ERNIE 2.0 [2], which is a continual pretraining transformer framework specializing in extracting the lexical, syntactic, and semantic information by augmenting knowledge training. The transformer module is composed of multi-head self-attention encoders.
- Two ERNIE units play different roles: one ERNIE 2.0 unit encodes global context, and another unit encodes local context. Their output is sent into the average pooling to reduce the dimension and obtain the generalized metaphor features.

WordNet Graph Embedding

- Graph embedding is an approach that is used to transform nodes, edges, and their features into vector space (a lower dimension) whilst maximally preserving properties like graph structure and information.
- The vector can represent some characteristics of the original node, such as if in the original structure between two nodes, then the two nodes as vector should be similar, just like word embedding. Because of that, we used graph embedding to introduce WordNet graph network information into the model processing with BiLSTM and reach some improvement.

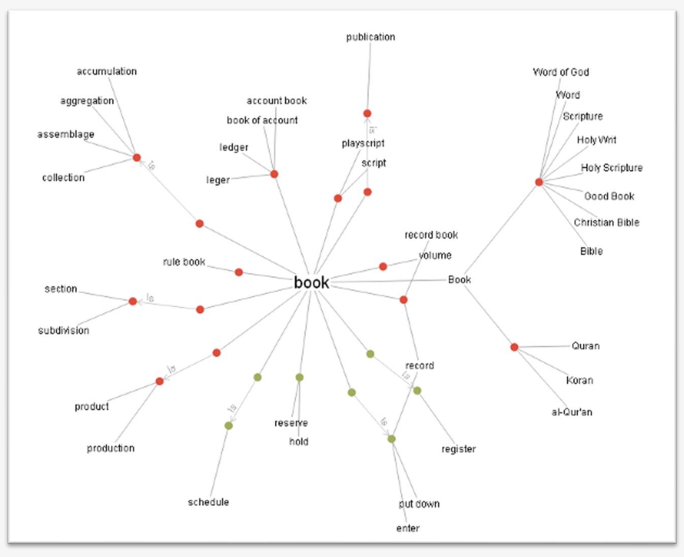

Fig 4: an example of WordNet graph

Experimental Result

- **Dataset**: VUA and MOH-X are used in the experiments.

  (1) VUA (VU Amsterdam Metaphor Corpus) is the largest manually labelled figurative language Corpus among various fields published in the metaphor detection task.

  (2) MOH-X corpus is a subset of the MOH corpus. The corpus texts are all from the WordNet with annotated for metaphorical verbs along with associated confidence scores.

- **Baseline**: A number of state-of-the-art neural network-based metaphor detection models have also implemented:

  (1) **Gao et al.(2018)** [3]: This model utilizes GloVe+ELMo for word embedding and BiLSTM as an encoder.

  (2) **RNN-HG** and **RNN-MHCA** [4]: These models utilize the BiLSTM+attention mechanism, published on ACL2019.

  (3) **BERT+MWE-Aware GCN** [5]: A metaphor detection model published on ACL2020 with BERT as encoder and Graph Convolution Network.

Model	VUA VERB			VUA ALL POS			MOH-X		
	P	R	F1	P	R	F1	P	R	F1
Gao et al. (2018)	53.4	65.6	58.9	68.2	71.3	69.7	75.3	84.3	79.1
RNN-HG	69.3	72.3	70.8	71.8	76.3	74.0	79.7	79.8	79.8
RNN-MHCA	66.3	**75.2**	70.5	**73.0**	75.7	**74.3**	77.5	83.1	80.0
BERT+MWE-Aware GCN	-	-	-	-	-	-	79.98	80.40	80.19
ERNIE+WordNet	**78.19**	73.71	**75.88**	66.84	**78.91**	72.37	**89.02**	**96.05**	**92.41**

Table 1: the results reveal our model reaches 75.88% F1 score on the VUA verb test set and 92.41% on the MOH-X dataset, outperforming other state-of-the-art metaphor detection models.

Conclusion and Future work

- We propose a novel framework with knowledge integration outcompeting each of the baselines to a great extent. The experiments show that our knowledge integration is a powerful tool for metaphor detection. In the future, we are going to utilize ensemble learning strategies to improve our model further.

References

[1] B. McCann, N. S. Keskar, C. Xiong, and R. Socher, "The natural language decathlon: Multitask learning as question answering,"arXivpreprintarXiv:1806.08730, 2018.

[2] Yu Sun, Shuohuan Wang, Yukun Li, Shikun Feng, HaoTian, Hua Wu, and Haifeng Wang. 2020. Ernie 2.0:A continual pre-training framework for language understanding. In Proceedings of the AAAI Conferenceon Artificial Intelligence, volume 34, pages 8968–8975.

[3] G. Gao, E. Choi, Y. Choi, and L. Zettlemoyer, "Neural metaphor detection incontext,"arXivpreprintarXiv:1808.09653, 2018

[4] R. Mao, C. Lin, and F. Guerin, "End-to-end sequential metaphor identification inspired by linguistic theories," in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 3888–3898, 2019.

[5] Omid Rohanian, Marek Rei, Shiva Taslimipoor, andLe An Ha. 2020. Verbal multiword expressions foridentification of metaphor. In Proceedings of the58th Annual Meeting of the Association for Computational Linguistics, pages 2890–2895, Online. Association for Computational Linguistics.