well probably due to multiple feature maps within each convolutional layer, where each feature map has the ability to capture certain semantic transition independently.

•**Query**: I was diagnosed with high blood pressure in my annual physical examination last month. What foods should I be eating on a regular basis? (上个月组织体检检查出了高血压，平时注意些啥该多吃什么?)
Prediction:

Rank	Intention	Probability
1	<disease,diet>	0.469885
2	<symptom,diet>	0.305881
3	<symptom,drug>	0.197734
4	<symptom,examine>	0.007157
5	<treatment,diagnose>	0.007089

In the query a user expresses his/her intention for seeking diet related information very implicitly in Chinese. Our model is able to put diet related intention at the first two places. Especially, the proposed model doesn't let "examine" related queries dominate its prediction simply due to the occurrence of word "examination" in the query. Actually "examine" in "physical examination" is mentioned as part of the background, which doesn't indicate "examine" related intentions.

•**Query**: How much does it costs for a Lumbar CT? Recently my lumbar always hurts. (腰椎CT检查大概需要多少费用？最近后腰老是酸疼。)
Prediction:

Rank	Intention	Probability
1	<examine,fee>	0.986955
2	<symptom,examine>	0.012433
3	<symptom,department>	0.000475
4	<disease,department>	8.50e-05
5	<examine,diagnose>	3.51e-05

Difficulty in detecting indention in this query is that there are tons of examinations one can receive in hospitals while we only have limited labeled queries which only cover a small portion of the examination terms. The proposed jointly modeling approach is feasible to solve this problem by integrating POS tag into the model. For example, if some other patient posts a query regarding the cost of cervical CT, not the lumbar in our case, then singular modeling methods such as NNID-FC only capture the fluctuation value in feature correlation matrix while the jointly modeling approach like NNID-JM can also utilize POS tags as long as the model is able to tag both "cervical CT" and "lumbar CT" with "n_examine". In that case, this special POS tag "n_examine" gives us extra information that some examination terms exist in the new query. The jointly modeling approach can take advantage of this extra clue to learn from previous knowledge more efficiently.

5. CONCLUSIONS

Intention detection for medical query will provide a new opportunity to connect patients with medical resources more seamlessly both in physical world and on the World Wide Web. Knowing what a user is looking for (e.g. a specific medicine that relieves headache, or a schedule of a doctor with high expertise in stomach diseases, or the average cost for lung surgery), health care resources can be made more accessible to the general public.

In this paper we present a jointly modeling approach to model intentions that users encoded in medical related text queries. By using feature-level modeling as one perspective to model intention, the proposed method is able to take variable-size text query as input. The resulting feature-level correlations are fed forward through an interleaving of convolution and pooling layer to extract semantic transitions from text queries. To aid the modeling, a bag-of-POS tags are used as the other perspective of modeling to indicate the existence of certain topic-specific words inside a query. Taking advantage of two heterogeneous perspectives in a joint model, the proposed approach has the capability to adapt to diverse user expressions covering a wide range of intentions. Performance evaluation and case studies have shown the effectiveness of the proposed method in discovering a complete and accurate intentions from text queries. Note that although in this work we focus on a medical related application, the proposed intention detection methods can be generalized and integrated into other existing applications such as recommendation system in search engines or phone call routing system for public services as well.

6. ACKNOWLEDGMENTS

This work is supported in part by NSF through grants III-1526499, CNS-1115234, and OISE-1129076. We thank Chaochun Liu, Tao Yang, Hao Wu and Yaliang Li who provided insights and expertise that assisted the research.

7. REFERENCES

[1] James F Allen and C Raymond Perrault. Analyzing intention in utterances. *Artificial intelligence*, 15(3):143-178, 1980.
[2] Yiming Yang and Xin Liu. A re-examination of text categorization methods. In *SIGIR*, 1999.
[3] Charu C Aggarwal and ChengXiang Zhai. *Mining text data*. Springer Science & Business Media, 2012.
[4] Bo Pang and Lillian Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1-135, 2008.
[5] Bing Liu and Lei Zhang. A survey of opinion mining and sentiment analysis. In *Mining text data*, pages 415-463. Springer, 2012.
[6] Yanshen Yin, Yong Zhang, Xiao Liu, Yan Zhang, Chunxiao Xing, and Hsinchun Chen. Healthqa: A chinese qa summary system for smart health. In *Smart Health*, pages 51-62. Springer, 2014.
[7] Edward Grefenstette and Phil Blunsom. A Convolutional Neural Network for Modelling Sentences. In *ACL*, 2014.
[8] Rie Johnson and Tong Zhang. Effective use of word order for text categorization with convolutional neural networks. In *NAACL-HLT*, 2015.
[9] Ronan Collobert, Jason Weston, Leon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural Language Processing (almost) from Scratch. *Journal of Machine Learning Research*, 12:2493–2537, 2011.

[10] Jian Hu, Gang Wang, Fred Lochovsky, Jian-tao Sun, and Zheng Chen. Understanding user's query intent with wikipedia. In *WWW*, 2009.

[11] Klaus-Peter Adlassnig. Fuzzy set theory in medical diagnosis. *Systems, Man and Cybernetics, IEEE Transactions on*, 16(2):260–265, 1986.

[12] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *NIPS*, 2013.

[13] Andriy Mnih and Geoffrey Hinton. Three new graphical models for statistical language modelling. In *ICML*, 2007.

[14] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *ICML*, 2008.

[15] Zheng Chen, Fan Lin, Huan Liu, Yin Liu, Wei-Ying Ma, and Liu Wenyin. User Intention Modeling in Web Applications Using Data Mining. In *WWW*, 2002.

[16] Botao Hu, Yuchen Zhang, Weizhu Chen, Gang Wang, and Qiang Yang. Characterizing search intent diversity into click models. In *WWW*, 2011.

[17] Honghua (Kathy) Dai, Lingzhi Zhao, Zaiqing Nie, Ji-Rong Wen, Lee Wang, and Ying Li. Detecting online commercial intention (OCI). In *WWW*, 2006.

[18] Puyang Xu and Ruhi Sarikaya. Convolutional neural network based triangular CRF for joint intent detection and slot filling. In *2013 IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU 2013 - Proceedings*, 2013.

[19] Long Chen, Dell Zhang, and Mark Levene. Question retrieval with user intent. In *SIGIR*, 2013.

[20] Xiao Ding, Ting Liu, Junwen Duan, and Jian-yun Nie. Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network. In *AAAI*, 2015.

[21] Puyang Xu and Ruhi Sarikaya. Contextual Domain Classification in Spoken Language Understanding Systems Using Recurrent Neural Network. In *ICASSP*, 2014.

[22] Xiang Zhang, Junbo Zhao, and Yann LeCun. Character-level convolutional networks for text classification. In *NIPS*, 2015.

[23] Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. Recurrent Convolutional Neural Networks for Text Classification. In *AAAI*, 2015.

[24] Ali S Razavian, Hossein Azizpour, Josephine Sullivan, and Stefan Carlsson. Cnn features off-the-shelf: an astounding baseline for recognition. In *CVPRW*, 2014.

[25] Yoon Kim. Convolutional Neural Networks for Sentence Classification. In *EMNLP*, 2014.

[26] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, 86(11):2278–2323, 1998.

[27] L Gershkoff-Stowe and S Goldin-Medow. Is there a natural order for expressing semantic relations? *Cognitive Psychology*, 45(3):375, 2002.

[28] Yulia Tsvetkov, Manaal Faruqui, Wang Ling, Guillaume Lample, and Chris Dyer. Evaluation of word vector representations by subspace alignment. In *EMNLP*, 2015.

[29] Hua-Ping Zhang, Hong-Kui Yu, De-Yi Xiong, and Qun Liu. HHMM-based chinese lexical analyzer ICTCLAS. In *SIGHAN*, 2003.

[30] Patrice Y Simard, Dave Steinkraus, and John C Platt. Best practices for convolutional neural networks applied to visual document analysis. In *ICDAR*, 2003.

[31] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580, 2012.

[32] Minjie Wang, Tianjun Xiao, Jianpeng Li, Jiaxing Zhang, Chuntao Hong, and Zheng Zhang. Minerva: A scalable and highly efficient training platform for deep learning. In *NIPS Workshop, Distributed Machine Learning and Matrix Computations*, 2014.

[33] Ilya Sutskever, James Martens, George Dahl, and Geoffrey E. Hinton. On the importance of initialization and momentum in deep learning. In *ICML*, 2013