Label Relation Graphs to Encode Prior Knowledge

Anshul Goyal
CSE, IIT Kanpur
anshulgo@iitk.ac.in

Md Enayat Ullah
MTH, IIT Kanpur
enayat@iitk.ac.in

Abheet Agarwal
MTH, IIT Kanpur
abheet@iitk.ac.in

1 Introduction

A crucial challenge in machine learning problems is that the rich structure of real world labels is generally not encoded in the traditional models. They either consider the label space overlapping (Logistic Regression) or mutually exclusive (Softmax Classifier). But in real-world scenarios, the label space is usually a mixture of both overlapping and mutually exclusive relations. There has been a couple of advances to incorporate prior knowledge in various problems. Vedantam et al made use of human generated abstract scenes made from clipart for learning semantic visual information [5]. Another line of work is NEIL which exploits the large-scale visual data to automatically extract common sense relationships [1]. This rich label structure is also useful in robot path planning wherein they extract information about objects from large amounts of natural language text [4]. An exciting approach is the use of Label relation graphs such as HEX (Hierarchical and Exclusion) graph to provide prior beliefs. Deng et al combined the HEX graph with a deep neural network which produced better results in Imagenet challenge and Zero-shot learning tasks [2].They further extended their approach to a probabilistic HEX model to subsume soft/uncertain relations between labels and formulated an algorithm to perform efficient exact inference on the graphical model [3].

2 Proposed Work

The first step is to study the approach and analysis of the original HEX paper as well as the recent probabilistic extension pHEX [2, 3]. We then intend to reproduce their implementation as well as empirically validate the model on other datasets. We plan to explore spatial relations between objects in images which is one of the possible directions suggested in the paper. In general, this framework is suitable for classification problems in which the label space exhibits a rich structure. Visual Question Answering is a task where prior knowledge could improve existing results. Our aim is to adapt the graphical framework for the same using rich label relations (from semantic dictionaries like Wordnet etc), and evaluate its performance.

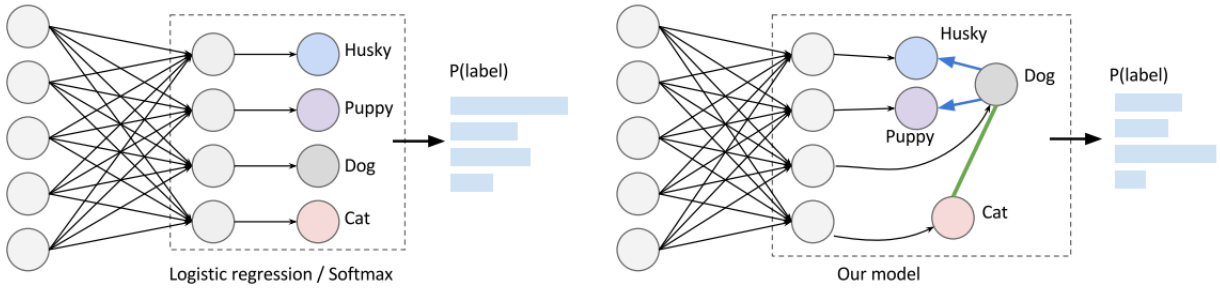

Figure 1: HEX Graph on Top of Deep Neural Network

Source: [2]

References

[1] X. Chen, A. Shrivastava, and A. Gupta. Neil: Extracting visual knowledge from web data. In *Computer Vision (ICCV), 2013 IEEE International Conference on*, pages 1409–1416. IEEE, 2013.

[2] J. Deng, N. Ding, Y. Jia, A. Frome, K. Murphy, S. Bengio, Y. Li, H. Neven, and H. Adam. Large-scale object classification using label relation graphs. In *Computer Vision–ECCV 2014*.

[3] N. Ding, J. Deng, K. Murphy, and H. Neven. Probabilistic label relation graphs with ising models. *arXiv preprint arXiv:1503.01428*, 2015.

[4] P. Kaiser, M. Lewis, R. P. Petrick, T. Asfour, M. Steedman, et al. Extracting common sense knowledge from text for robot planning. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE.

[5] R. Vedantam, X. Lin, T. Batra, C. Lawrence Zitnick, and D. Parikh. Learning common sense through visual abstraction. In *Proceedings of the IEEE International Conference on Computer Vision*.