M-Protain Diagnostics Using Generative Active Learning

Hanyu Li
Department of Physics
Tsinghua University
l-hy16@mails.tsinghua.edu.cn

Junsong Yuan
Department of Computer Science and Engineering
University at Buffalo
jsyuan@buffalo.edu

Abstract

With comparatively less labeled data and high labeling cost, most of the medical involved tasks can not be directly tackled by state of art machine learning approaches for their demand of large carefully labeled datasets.

Introduction

As deep models achieve astonishing results in almost every machine learning tasks, some unavoidable problems such as the need for large carefully labeled dataset has troubled researchers from the start. Part of the reason behind the tremendous success in deep learning is the availability of large-scale labeled data(Sun et al. 2017). Although data labeling companies and platforms claim that they can provide inexpensive yet high quality data(Buhrmester, Kwang, and Gosling 2011), achieving such datasets can be extremely costly or even unrealistic in the scenarios where labeling requires high professionality. For instance, some medical image tasks can not be labeled by people without systematic training. However, the small number of these experts has determined that large-scale dataset is difficult to biuld. Plus, they are probably already preoccupied.

M-protain can be detected by immunofixation electrophoresis(IFE) images, and different categories of results may indicate MGUS, multiple myeloma, etc.

Active learning and GAN

Our paper try to

Related Work

Multiple techniques were used to tackle with small datasets and partially labeled datasets including active learning(Settles 2009), generative models(Goodfellow et al. 2014)(Kingma and Welling 2013), data augmentation(Tanner and Wong 1987), and domain transfer(Pan and Yang 2009) etc. asdlfdklflkdjflkadf;asdafsdjklfajfkjadf(Zhu and Bento 2017)

Preliminaries

The problem is defined as below

Proposed Method

Our method make use of

Experiments

I did such experiments

Discussion

After I did these experiments

Conclusion

To sum up

References

Buhrmester, M.; Kwang, T.; and Gosling, S. D. 2011. Amazon's mechanical turk: A new source of inexpensive, yet high-quality, data? *Perspectives on psychological science* 6(1):3–5.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial nets. In *Advances in neural information processing systems*, 2672–2680.

Kingma, D. P., and Welling, M. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

Pan, S. J., and Yang, Q. 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22(10):1345–1359.

Settles, B. 2009. Active learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences.

Sun, C.; Shrivastava, A.; Singh, S.; and Gupta, A. 2017. Revisiting unreasonable effectiveness of data in deep learning era. In *Proceedings of the IEEE international conference on computer vision*, 843–852.

Tanner, M. A., and Wong, W. H. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association* 82(398):528–540.

Zhu, J.-J., and Bento, J. 2017. Generative adversarial active learning. *ArXiv* abs/1702.07956.