Paper Review for the Computer Science Research Course and the Second Exam

Xiaoke(Jimmy) Shen

September 22, 2017

1 Introduction

This paper review report is provided based on the requirement of the computer science research course provided by the Graduate Center of the City University of New York during the 2017 Fall Semester. The main objective of this course is helping the PHD candidates to prepare their second exam and identify their thesis topic as early as possible [1].
In this report the papers related to the object classification, object detection and object semantic segmentation for the 2D and 3D objects will be discussed. As most of the state of the art algorithms used for those tasks are based on the deep convolutional neural networks, the important papers related to the deep neural networks will also be discussed in this article.

The structure of this article is described bellow: The papers related to the theory part of the deep learning will be discussed in the second session. In the third session, the papers in the 2D object classification will be discussed. In session 4, the 2D object detection papers will be studied. At the same time, an interesting and more challenge related to the 2D image procession or computer vision will be discussed in session 5 which is the 2D object semantic segmentation. In the rest part of this article, the similar tasks in 3D will be discussed as the final goal of this paper review is finding some possible approaches to improve the current 3D computer vision algorithms based on the state of the are 2D computer vision algorithms.

2 Deep Learning Theory

2.1 Approximation with Artificial Neural Networks [2]

In order to build the mathematical theory of the artificial neural networks, several papers are published in the 20 century. In this paper one main contribution is the universal approximation theorem with proof. The universal approximation theorem claims [2] that the standard multilayer feed-forward networks with a single hidden layer that contains finite number of hidden neurons, and with arbitrary activation function are universal approximators in $C(R^m)$. The universal approximation theorem is one of the important theoretical support for the artificial neural networks. However, at that time as the huge size labeled data sets are not available, the updated algorithms haven't been invented and also the limited computation power, these ideas can not be verified.

2.2 Approximation Capabilities of Multilayer Feedforward Networks [3]

The unique value of the Kurt Hornik (1991) [3] paper is it showed that it is not the specific choice of the activation function, but rather the multilayer feedforward architecture itself which gives neural networks the potential of being universal approximators. This is an important contribution which is the foundation for the current state of the art deep learning architecture such as VGG 16 [4] and resnet [5]

[6] [7] [8] [9] [10]

[12] [**?**]

[13] [14]

2.3 Learning representations by back-propagating errors [15]

The paper of 1986 significantly contributed to the popularisation of BP(Back Propagation) for NNs [15], experimentally demonstrating the emergence of useful internal representations in hidden layers.

This paper will be read in the future.

3 Object Classification for 2D Images

3.1 Backpropagation applied to handwritten zip code recognition [17]

The first important application of using the BP(Back Propagation) to well resolve the real life problem from the literature. From this paper, one important structure of the neural network including the layers with filters were introduced. The similar structure is used in the modern neural network structures such as Alex Net [11], VGG 16 [4] and resnet [5]

4 Object Detection for 2D Images

References

[1] P. Ji, "Syllabus of the computer science research at cuny fall 2017," 2017.

[2] B. C. Csji, "Approximation with artificial neural networks," pp. 11–12, 2001.

[3] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. vol. 4, 1991.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015.

[6] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *CoRR*, vol. abs/1311.2524, 2013.

[7] R. B. Girshick, "Fast R-CNN," in *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pp. 1440–1448, IEEE Computer Society, 2015.

[8] S. Ren, K. He, R. B. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015,*

Montreal, Quebec, Canada (C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, eds.), pp. 91–99, 2015.

[9] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015.

[10] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *CoRR*, vol. abs/1612.08242, 2016.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25* (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, Curran Associates, Inc., 2012.

[12] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.

[13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *CoRR*, vol. abs/1502.03167, 2015.

[14] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: common objects in context," *CoRR*, vol. abs/1405.0312, 2014.

[15] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 10 1986.

[16] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85 – 117, 2015.

[17] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541–551, 1989.