Progress Report

Bardia Mojra

March 13, 2020

Robotic Vision Lab
The University of Texas at Arlington

1 Progress

Following items are listed in order of priority:

- Based on what I learned from [1], [2], [3], [4], and image super resolution work with CNNs, I agree that the combination of two methods, Deep Learning (DL) and Reinforcement Learning (RL), would create a more powerful machine learning method. CNNs provide a relatively efficient way of processing high resolution frames and DL provides models for learning with strategy, with long term goal in mind. Since I am somewhat familiar with CNNs, I decided to cover basics of RL. Without some basic understanding of RL, these papers will not make much sense. I will continue to read on the topic for few more days and then begin working on DISCOMAN [5] presentation.
- Since a few weeks ago, I began using python for small tasks and to enhance C code debugging. Now, I am comfortable with the syntax and ready to use more advance libraries.
- Read [2], began to dissect it. It's a heavy paper, I have already read it twice, I need to go back and read it again and learn the material.
- Read [3], still need to dissect it.

2 Plans

Following items are listed in order of priority:

- (Top priority)I will begin dissecting [5] to prepare for March 27th presentation.
- (On pause) I am working through Jason Brownlee's ML Mastery book, [6].
- (On pause) Resume Robotic Perception course as soon as possible.
- (On pause) Resume Machine Learning course with Andrew Ng as soon as possible.
- (On pause) Need to read [7], [8], [9], and [10]; these papers seem fundamental to understanding the overall picture.

- (On pause) Get intimate with Python, Numpy, Pandas, Scipy, and Matplotlib, TensorFlow and PyTorch.
- (On pause) Read Digital Image Processing by Gonzalez and Woods.
- (On pause) Learn ROS.
- (On pause) (Supremely important) Read on scene understanding, semantic SLAM, graph SLAM, visual odometer, place recognition, and Kalman Filtering. Read Niko Sunderhauf's research publications.

3 General Notes

This section summarizes general research leads. The following item are to be investigated, understood and briefly summarized.

3.1 Keywords

- Bayesian Learning: This probabilistic ML approach treats model parameters as random variables. Read [11] for more details.
- Convex Optimization:
- Q-Learning: A learning model of reinforcement (RL)
- Deep Reinforcement Learning:

4 Literature Review

4.1 Dynamic Graph CNN for Learning on Point Clouds [2]

This paper introduces a new model for training CNNs to learn similar features of point cloud objects.

4.1.1 Keywords

- PointNet
- Extrinsic and intrinsic descriptors:
- Permutation variance:

4.2 Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning [3]

4.3 Single Image Super-Resolution Using Multi-Scale Convolutional Neural Network - MSSR [12]

Paper proposes an architecture with two parallel path with different depths (which correspond to scales) for residual learning; where one path (module L) is used for large factor up-scaling (x4, x8) and the other (module S) for small factor up-scaling (x2). At the end, it combines the outputs by summation (a form of ensembles). In contrast to previous work where the focus is on small factor up-scaling (x2) and repeat if needed, this model takes higher factor up-scaling into consideration while training the network which helps with reducing blurriness of output image for higher factor up-scaling. The model uses multi-scale residual learning to train on general model for multiple up-scaling factors; hence, saving memory and processing time. This paper provides experiment results that show higher output image integrity where peak signal to noise ratio (PSNR) and structural similarity index (SSIM) are higher or comparatively close to state-of-the-art methods.

4.3.1 Keywords

- Bicubic interpolation: A method for artificially up-sampling an image, it is superior to nearest neighbor and bi-linear interpolation. Rather than simply averaging neighboring pixels to calculate pixel value in between, it also takes into account the gradient of its 16 neighboring pixels (hence bicubic) and results in a smoother curved image. This happens to work better than previous methods because images of physical world are made up of mostly round edges rather than sharp and square ones.
- Lanczos re-sampling:
- Statistical priors:
- Stochastic Neighbor Embedding (paper by G. Hinton [4]): Read paper.
- Parse coding:
- GoogLeNet:
- YCbCr color space:

- Caffe package (paper by Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.): Convolutional architecture for fast feature embedding. Read and write literature review.

- Adam method (paper by Kingma, D., Ba, J.): A method for stochastic optimization.

- PSNR (performance metric): Peak Signal to Noise Ratio represents the ratio between max (peak) possible value (power) of a signal (image) and the power of distorting noise.

- SSIM (performance metric): Structural Similarity Index, is the ratio of structural features of a processed image to the original image. The value represent percentage of structural features/information retained throughout image processing.

- A+ (SR method, paper):

- SelfEx (SR method, paper):

- SRCNN (SR method, paper): Image Super Resolution Using Deep Convolutional Networks. Read again and write literature review.

- FSRCNN (SR method, paper):

- VDSR (SR method, paper):

4.4 Value Iteration Networks - VIN [1]

4.4.1 Keywords

- Imitation Learning: **need to finish reading [13].**

- CNNs applied to reinforcement learning:

- MDP (Markov Decision Process):

- VI Algorithm (Value Iteration):

- SGD (stochastic gradient decent):

- Theano Code: "Theano is a Python library that allows you to define, optimize, and evaluate mathematical expressions involving multi-dimensional arrays efficiently. It can use GPUs and perform efficient symbolic differentiation."

- Grid-World Domain:

References

[1] A. Tamar, Y. WU, G. Thomas, S. Levine, and P. Abbeel, "Value iteration networks," in *Advances in Neural Information Processing Systems 29* (D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, eds.), pp. 2154–2162, Curran Associates, Inc., 2016.

[2] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph CNN for learning on point clouds," *CoRR*, vol. abs/1801.07829, 2018.

[3] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*, ICML'16, p. 1050–1059, JMLR.org, 2016.

[4] G. Hinton and S. Roweis, "Stochastic neighbor embedding," *Advances in neural information processing systems*, vol. 15, pp. 833–840, 2003.

[5] P. Kirsanov, A. Gaskarov, F. Konokhov, K. Sofiiuk, A. Vorontsova, I. Slinko, D. Zhukov, S. Bykov, O. Barinova, and A. Konushin, "Discoman: Dataset of indoor scenes for odometry, mapping and navigation," 2019.

[6] J. Brownlee, *Machine Learning Mastery With Python: Understand Your Data, Create Accurate Models, and Work Projects End-to-End.* Machine Learning Mastery, 2016.

[7] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *CoRR*, vol. abs/1501.00092, 2015.

[8] D. Liu, Z. Wang, N. M. Nasrabadi, and T. S. Huang, "Learning a mixture of deep networks for single image super-resolution," *CoRR*, vol. abs/1701.00823, 2017.

[9] K. Doherty, D. Fourie, and J. Leonard, "Multimodal semantic slam with probabilistic data association," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 2419–2425, May 2019.

[10] F. Wang, M. Jiang, C. Qian, S. Yang, C. Li, H. Zhang, X. Wang, and X. Tang, "Residual attention network for image classification," *CoRR*, vol. abs/1704.06904, 2017.

[11] “Bayesian learning,” 2012.

[12] X. Jia, X. Xu, B. Cai, and K. Guo, “Single image super-resolution using multi-scale convolutional neural network,” *Lecture Notes in Computer Science*, p. 149–157, 2018.

[13] V. Kurin, “Introduction to imitation learning,” Aug 2017.