**Refining Deep Reinforcement learning in Image Restoration**

E6885\_2018Fall\_report

Zhi Ji zj4226, Boyuan Sun bs3113

*Columbia University*

**Abstract**

*We introduce our approach to refine deep reinforcement learning used in image restoration. The image restoration network uses a tool box of many small scale convolutional neural networks for very specific tasks instead of using a large scale deep convolutional neural train on a wide range of samples and different levels of severities. The agent needs to choose the appropriate tool of the right type and severity of the distortion. This process can be done using Deep Q-learning Network. We provide multiple approaches to improve the DQN. The methods we represent combining three main perspectives: Double Deep Q-learning Networks, Dueling Deep Q-learning Network and prioritized experience replay. All three methods have been formulated and combined to the image restoration network. Their potentials and advantages are studied and explored****.***

**Introduction**

Introduced since 2014, the deep Q-learning networks (DQN) have been widely apply to solve Markov Decision Processes (MDPs) for the optimal solution. Since then many improvements has been made including double DQN, dueling deep Q-network (DDQN), and prioritized experience replay (PER). We introduce our approach to this image restoring task using a reinforcement learning based CNN tool selection LSTM network with three improvements in deep Q-learning. Firstly, Deep Q-network is replaced by the Double DQNs. Double DQNs provide our network robustness towards overestimations of action values caused by traditional maximum action value estimation in DQN.

Secondly, dueling DQN is added to our network architecture. While the output of original DQN is the Q value of each action taken by agent, dueling DQN output both state values and state advantages of each actions. However, this method does not guarantee the uniqueness of the state values and state advantage for single action we take. To further improve our dueling DQN, we designed to replace the state advantage value with single averaged advantage value to increase the stability and ensure that is the single action solution for greedy policy.

Lastly, we add prioritized replay to our DQN. Instead of updating the batch stochastically, we introduced the TD-error to our system and update batches based on the value of TD-error. The larger the TD-error of a batch is, the more likely the prediction of that batch can be improved. Priorities of each states are also being stored in a tree structure for easy search and update.

**Background/Related Work**

Image restoration has been well studied in the field of deep learning using convolutional neural networks. Images that are corrupted by noise and JPEG artifacts can be restored to a high-quality output. Many studies present a single, deep complex, human made network trained for certain tasks. Tasks include deblurring [31, 35, 42], denoising [6, 24], JPEG artifacts reduction [7, 9, 41] and super-resolution [8, 17, 19, 20, 22, 36, 37, 39]. Our project works on another route to the similar problem.

Using a deep Q-learning Network, we are able to achieve the same or better result for various tasks comparing to studies mentioned above. With addition of three different ways to improve the deep reinforcement learning network, we accomplish the network performance on the next level providing with extra robustness and speed to our architecture.

There are many works that have been done prior to ours and many of them use deep CNN to handle image restoration related problems. Kim et al. [19] developed a 20-lay CNN network to solve multi-scale image super-resolution tasks. Zhang et al. [44] used a 20-layer deep CNN to deal with multiple restoration tasks by multi-threading. In addition, one of them innovatively used Deep Q-learning Network [1] that successfully reduced the complexity of deep networks in prior works to a small scale with multiple shallower CNNs.

Our works were developed on the top of [1] and introduced three improvements: Double Q-learning, Prioritized Experience Replay[2] and Dueling Q-learning.

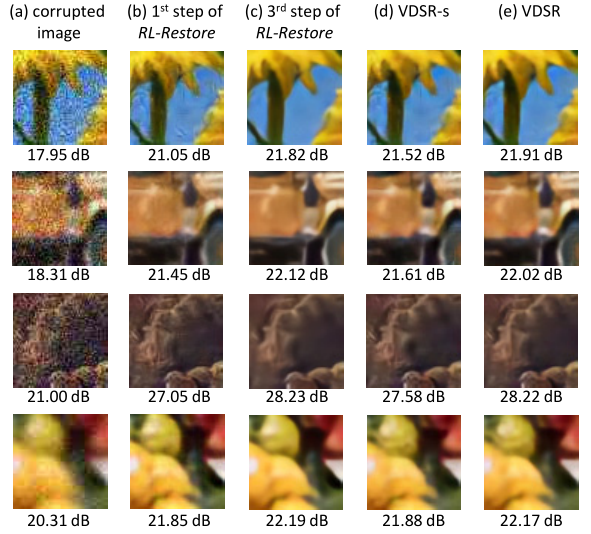
****

Figure 1. (a) images corrupted by various distortions. (b-c) Each step a specific tool is selected by the agent to improve the image quality. (d-e) CNN-based results.

**Approach: algorithm development/theoretic results**

0. To restore a clear image I1 from a given distorted image I2 to the ground truth image I3. The toolchain shown in the figure 3 is introduced. At each step t, the agent observes the current state St , the current restored image I2 and the output of the agent at the previous step vt. Note that I1 represents the input image and v1 is a zero vector. Based on the maximum value of the agent’s output vt , an action a t is selected and the corresponding tool is used to restore the current image. Steps of restoration iteratively replays until the stopping action is selected.

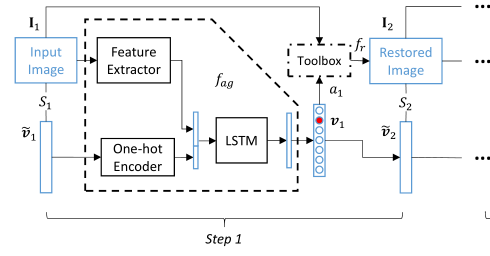


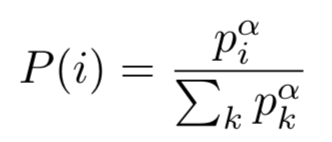
Figure 3, a single iteration of toolchain selection to restore one kind of distortion.

1. Double Deep Q-learning Network

In the Deep Q-learning Network, we use double q network structure, q-prediction network and q-target network. The q-prediction network is updated by the q-target network each step. In another word, the q-target network is frozen in time. Since the data structure that have been stored in the memory is (s,a,r,s') and the q-target is calculated base on the reward R and maximum output Q(s',a') that obtained by inputting s' to our q-target network. During this process, the selection of action a' using s' and predict Q(s',a') are using the same Q value. As a result, overestimation is expected to occur in this case. To avoid the overestimation, Double DQN update the action a' using Q value from another architecture and this could balance the overestimation brought by using the same Q value.

2. Prioritized Experience Replay

Beside the q-target and q-prediction network we mentioned above in our approach one, memory replay also plays an important role in DQN. It solves the relativity and dynamic distribution problem during reinforcement learning process. It stores the transaction sample (st,at,rt,st+1) to memory replay storage in minibatch. For small reward causing slow training speed, stochastic replay usually can not generate an efficient transition in humongous errors trials. To deal with this problem, we sample batch based on their priorities calculated using TD-error. We calculate the priority as

 (1)

To efficiently sample from distribution (1), the complexity cannot depend on N. We save and sample the memory through sum-tree [2] data structure as shown in Figure 4.

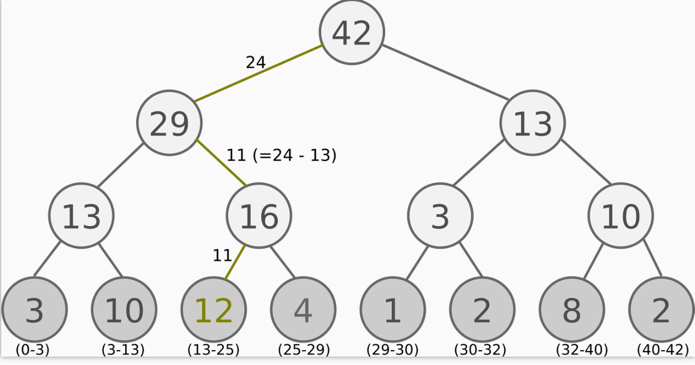


Figure 4, a single example of SumTree structure where every node is the sum of its children.

3. Dueling Deep Q-learning Network

In DQN, the output of network is the q-value of each actions. On the other hand, Dueling DQN outputs state value combined with the advantage of each action in this state to generate the q-value.

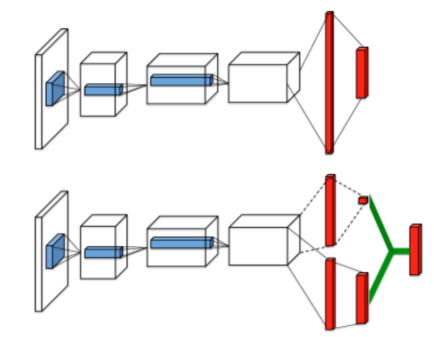


Figure 5, demonstration of dueling DQN

**Experiment results**

**Conclusion**

We have successfully improved the performance of DQN in image restore using Deep Q-learning Network. As  
mentioned in abstract three improvements, double DQN, dueling deep Q-network, and prioritized experience replay, have been successfully implemented. Our approach greatly improved the PNSR as well as training speed.

**References**

[1] K. Yu, C. Dong, L. Lin, C. Change Loy. Crafting a Toolchain for Image Restoration by Deep Reinforcement Learning. In CVPR, 2018.

[2] T. Schaul, J. Quan, I. Antonoglou and D. Silver. Prioritized Experience Replay. In ICLR, 2016.

[3] B. Baker, O. Gupta, N. Naik, and R. Raskar. Designing neural network architectures using reinforcement learning. In ICLR, 2017.

[4] Q. Cao, L. Lin, Y. Shi, X. Liang, and G. Li. Attention-aware

Face hallucination via deep reinforcement learning. InCVPR,

2017.

[5] W. Chen, J. Wilson, S. Tyree, K. Weinberger, and Y. Chen.

Compressing neural networks with the hashing trick. In

ICML, 2015.

[6] Y. Chen, W. Yu, and T. Pock. On learning optimized reaction

diffusion processes for effective image restoration. In CVPR,

2015.

[7] C. Dong, Y. Deng, C. C. Loy, and X. Tang. Compression ar-

tifacts reduction by a deep convolutional network. In ICCV,

2015.

[8] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-

resolution using deep convolutional networks. TPAMI,

38(2):295–307, 2016.

[9] J. Guo and H. Chao. Building dual-domain representations

for compression artifacts reduction. In ECCV, 2016.

[10] J. Guo and H. Chao. One-to-many network for visually

pleasing compression artifacts reduction. arXiv preprint

arXiv:1611.04994, 2016.

[11] S. Han, H. Mao, and W. J. Dally. Deep compression: Com-

pressing deep neural networks with pruning, trained quanti-

zation and huffman coding. In ICLR, 2016.

[12] M. Hausknecht and P. Stone. Deep recurrent q-learning for

partially observable mdps. CoRR, abs/1507.06527, 2015.

[13] Y. He, K. Cao, C. Li, and C. C. Loy. Merge or not? learning

to group faces via imitation learning. In AAAI, 2018.

[14] G. Hinton, O. Vinyals, and J. Dean. Distilling the knowledge

in a neural network. In NIPS Workshop, 2014.

[15] S. Hochreiter and J. Schmidhuber. Long short-term memory.

Neural Computation, 9(8):1735–1780, 1997.

[16] C. Huang, S. Lucey, and D. Ramanan. Learning policies for

adaptive tracking with deep feature cascades. In ICCV, 2017.

[17] T.-W. Hui, C. C. Loy, and X. Tang. Depth map super-

resolution by deep multi-scale guidance. In ECCV, 2016.

[18] J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for

real-time style transfer and super-resolution. In ECCV, 2016.

[19] J. Kim, J. Kwon Lee, and K. Mu Lee. Accurate image super-

resolutionusingverydeepconvolutionalnetworks. InCVPR,

2016.

[20] J. Kim, J. Kwon Lee, and K. Mu Lee. Deeply-recursive

convolutional network for image super-resolution. In CVPR,

2016.

[21] D. Kingma and J. Ba. Adam: A method for stochastic opti-

mization. In ICLR, 2015.

[22] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep

laplacian pyramid networks for fast and accurate super-

resolution. In CVPR, 2017.

[23] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham,

A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al.

Photo-realistic single image super-resolution using a genera-

tive adversarial network. In CVPR, 2017.

[24] S. Lefkimmiatis. Non-local color image denoising with con-

volutional neural networks. In CVPR, 2017.

[25] X. Liang, L. Lee, and E. P. Xing. Deep variation-structured

reinforcement learning for visual relationship and attribute

detection. In CVPR, 2017.

[26] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez,

Y. Tassa, D. Silver, and D. Wierstra. Continuous control with

deep reinforcement learning. In ICLR, 2016.

[27] L.-J. Lin. Reinforcement learning for robots using neural

networks. PhD thesis, Fujitsu Laboratories Ltd, 1993.

[28] B. Liu and X. He. Learning dynamic hierarchical models for

anytime scene labeling. In ECCV, 2016.

[29] V. Mnih, N. Heess, A. Graves, et al. Recurrent models of

visual attention. In NIPS, 2014.

[30] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness,

M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland,

G. Ostrovski, et al. Human-level control through deep rein-

forcement learning. Nature, 518(7540):529–533, 2015.

[31] S. Nah, T. H. Kim, and K. M. Lee. Deep multi-scale con-

volutional neural network for dynamic scene deblurring. In

CVPR, 2017.

[32] Z. Ren, X. Wang, N. Zhang, X. Lv, and L.-J. Li. Deep reinforcement learning-based image captioning with embedding

reward. In CVPR, 2017.

[33] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre,

G. Van Den Driessche, J. Schrittwieser, I. Antonoglou,

V. Panneershelvam, M. Lanctot, et al. Mastering the game

of Go with deep neural networks and tree search. Nature,

529(7587):484–489, 2016.

[34] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou,

A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton,

et al. Mastering the game of Go without human knowledge.

Nature, 550(7676):354–359, 2017.

[35] J. Sun, W. Cao, Z. Xu, and J. Ponce. Learning a convolu-

tional neural network for non-uniform motion blur removal.

In CVPR, 2015.

[36] Y. Tai, J. Yang, and X. Liu. Image super-resolution via deep

recursive residual network. In CVPR, 2017.

[37] Y. Tai, J. Yang, X. Liu, and C. Xu. Memnet: A persistent

memory network for image restoration. In ICCV, 2017.

[38] J. Vermorel and M. Mohri. Multi-armed bandit algorithms

and empirical evaluation. In ECML, 2005.

[39] X. Wang, K. Yu, C. Dong, and C. C. Loy. Recovering realis-

tic texture in image super-resolution by deep spatial feature

transform. In CVPR, 2018.

[40] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli.

Image quality assessment: from error visibility to structural

similarity. TIP, 13(4):600–612, 2004.

[41] Z. Wang, D. Liu, S. Chang, Q. Ling, Y. Yang, and T. S.

Huang. D3: Deep dual-domain based fast restoration of

JPEG-compressed images. In CVPR, 2016.

[42] L. Xu, X. Tao, and J. Jia. Inverse kernels for fast spatial

deconvolution. In ECCV, 2014.

9

[43] S. Yun, J. Choi, Y. Yoo, K. Yun, and J. Y. Choi. Action-

decision networks for visual tracking with deep reinforce-

ment learning. In CVPR, 2017.

[44] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond

a gaussian denoiser: Residual learning of deep cnn for image

denoising. TIP, 2017.

[45] B. Zoph and Q. V. Le. Neural architecture search with rein-

forcement learning. In ICLR, 2017.

Contributions of Each Member

|  |  |  |
| --- | --- | --- |
|  | Experiment | Code |
|  |  |  |
|  |  |  |