Exploration for basic binary convolution unit

Wenxuan Zhang
Wzhang121@hawk.iit.edu

ABSTRACT

To be completed..

ACM Reference Format:
Wenxuan Zhang. 2024. Exploration for basic binary convolution unit. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With multimedia and the internet becoming an integral part of everyday life, and the increasing prevalence of low-quality images due to constraints such as shooting environments and equipment, tasks such as removing artifacts from lossy compressed images and restoring high-quality images have become essential branches of computer vision. Image restoration (IR) aims to recover high-quality images from low-quality ones to meet various application needs[1]. Image restoration techniques are primarily aimed at addressing the "degradation" in the imaging process, which refers to various factors affecting the imaging system, such as defocusing of the imaging system, relative motion between equipment and objects, or inherent defects in equipment, leading to images not meeting ideal quality requirements. The purpose of image restoration techniques is to use deblurring functions to remove blurry parts from images and restore the true nature of the image to achieve the goal of repairing images.

Traditional image restoration tasks include image denoising, super-resolution, and reducing compression artifacts, all of which are crucial for enhancing image quality and ensuring accurate visual perception[12]. From an application perspective, common image restoration methods can be roughly divided into two main categories: frequency domain restoration algorithms[4] and spatial domain iterative algorithms[8]. Frequency domain restoration algorithms directly perform restoration calculations on images in the frequency domain using restoration filters[4]. These algorithms analyze and process the frequency spectrum of images to achieve objectives such as noise removal and detail enhancement. On the other hand, spatial domain iterative algorithms solve restoration problems using iterative methods[8]. Since the restoration process cannot be described by a linear system, restoration is inherently nonlinear and requires iterative approaches to gradually approach the optimal solution. However, image restoration tasks often face challenges related to ill-posedness and computational complexity. Ill-posedness refers to the presence of multiple solutions to the problem itself, while computational complexity involves handling large amounts of data and complex mathematical operations. Therefore, reducing computational and memory resource consumption while ensuring algorithm performance has become an important challenge. Overall, traditional image restoration tasks are critical for various applications in computer vision and image processing, and addressing the challenges associated with these tasks requires innovative solutions and optimization techniques.

In recent years, with the rise of deep learning, neural network-based image restoration methods have gradually become a research hotspot. These methods leverage the powerful fitting and feature extraction capabilities of deep neural networks to effectively address image restoration problems and, to some extent, overcome some limitations of traditional methods. Among them, CNNs and GANs have made significant progress in image restoration technology. CNNs primarily approximate the mapping function from input images to output images through angle learning[11]. They demonstrate strong prior modeling capabilities within the deep learning framework and leverage the parallel computing power of GPUs, leading to significant advancements in image restoration

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

techniques in recent years. On the other hand, GANs utilize generators to produce realistic images and discriminators to distinguish between generated and real images. Through adversarial training, generators and discriminators continuously optimize their abilities, enabling generators to produce realistic images and thus achieve the goal of image restoration[9].

However, deep learning methods also face challenges such as large data requirements, long training and inference times, and increasing demands for computational resources. To reduce computational costs, researchers have begun to focus on binaryizing images, based on binary convolutional models. Binary Neural Networks (BNNs) transform the pixel units of images into a binary form containing only -1s and 1s, used to represent weights and activation functions[10]. Compared to full-precision neural networks, they can perform convolution operations using simple combinations like XNOR+popcount instead of float32 multiplications and accumulations, thus saving a significant amount of memory and computation, greatly facilitating model deployment on resource-constrained devices[7]. However, due to the limited information that binary representation can convey, BNNs have consistently exhibited lower model accuracy compared to full-precision models. Although recent research efforts such as MeliusNet[2], IRNet[5], and ReActNet[6] have managed to raise the Top-1 accuracy of BNNs on the ImageNet dataset to above 0.70, reducing the gap to about 3 percentage points from the corresponding full-precision models, they have also increased computational complexity.

This paper comprehensively explores the impact of binary networks on restoration networks. After extensive experimentation and analysis, it proposes a basic binary convolution unit (BBCU) and binaryization scheme for image restoration. Compared to previous binary restoration networks, significant performance improvements have been achieved in tasks such as super-resolution, denoising, and compression artifact reduction. For instance, in the super-resolution task, it achieves up to a 10x speedup while maintaining the same performance[3].

Compared to full-precision (32-bit) CNN models, Binary Neural Networks (BNNs) can achieve up to 32x memory savings. Additionally, since activations are also binaryized, binary convolution operations can be implemented using bitwise XNOR and popcount operations. Unlike full-precision convolutions that require floating-point operations, binary convolutions rely on bitwise operations, leading to up to 64x computational savings.

Binary neural networks are considered one of the most promising methods for neural network compression, leading to numerous studies based on BNNs, albeit mostly focusing on advanced image processing techniques such as classification. This paper explores the application of BNNs in low-level visual tasks like image denoising, conducting experiments and analyzing residual connections, batch normalization (BN), and activation functions[3], upon which a robust Basic Binary Convolutional Unit (BBCU) is designed. In image restoration tasks, residual connections are found to be crucial for binarized image restoration networks. Since BNNs binarize the input full-precision activations into 1 or -1 before binary convolution, it implies that BNNs lose a significant amount of information regarding the activation value range. By adding full-precision residual connections to each binary convolution, BNNs can mitigate the loss of value range information. Analysis reveals that batch normalization in BBCU plays a beneficial role in image restoration tasks by helping balance residual connections and binary convolutions. Building upon these findings, researchers propose a residual alignment (RA) scheme to enhance performance by increasing the value range of residual connections and simplifying BNN structures[3]. Additionally, it is observed during experiments that moving activation functions to residual connections can further improve performance. Experimental results demonstrate significant performance improvements of BBCU across multiple restoration tasks.

2 DATASET

2.1 DIV2K

DIV2K is an open dataset created by Radu Timofte and Eirikur Agustsson in 2017 for Super-Resolution research. The full name of this dataset is "DIVerse 2K resolution dataset," comprising approximately 800 images sourced from various origins including films, natural landscapes, animations, and computer-generated imagery, among others. These images are high-resolution, with a resolution of 2048x1024 pixels, hence the term "2K". Additionally, it provides high-quality low-resolution versions, obtained by downsampling the high-resolution images, as well as high-resolution label images used for evaluating the resolution algorithm's performance.

This dataset contains 1000 low-resolution images with different degradation types, divided into: • Training data: 800 low-resolution images, along with high-resolution and low-resolution images for degradation factors. • Validation data: 100 high-resolution images used to generate low-resolution images. • Test data: 100 diverse images used to generate low-resolution images.

The DIV2K dataset is primarily compiled from various publicly available high-resolution image resources, such as photography websites and online image libraries. It selects images with high clarity and rich details to ensure dataset quality. The selected images undergo preprocessing, including cropping, scaling, and color space conversion, to adapt to the requirements of super-resolution tasks. Different degradation models, such as bicubic downsampling, blur, and noise addition, are applied to the high-resolution images to generate corresponding low-resolution images. The purpose of the DIV2K dataset is to provide a broad and challenging dataset to promote research and evaluation of image super-resolution algorithms. These images encompass various complex textures, structures, and contents, covering different scenes and objects, thereby better representing images in the real world. Due to its extensive coverage and high-quality image content, the DIV2K dataset has become one of the widely used benchmark datasets in the field of image super-resolution research. In this study, all models are trained on the DIV2K dataset to evaluate the performance of algorithms in improving image resolution and compare and compete with other algorithms.

Despite being highly popular in the field of super-resolution, the DIV2K dataset also has some potential limitations. For example, there may be diversity issues as the dataset primarily focuses on super-resolution of natural images, which may limit the model's generalization ability to specific scenes or objects. Moreover, the low-resolution images in the DIV2K dataset are generated using simple degradation models, which may not fully simulate the complex degradation processes in the real world, such as camera shake and motion blur. Models trained on DIV2K may not achieve the desired results in practical applications due to different degradation conditions arising from various types of devices and settings.

In conclusion, while the DIV2K dataset serves as an important resource in the field of super-resolution, providing researchers with a large amount of high-quality training data, considerations need to be made regarding the dataset's diversity and the complexity of degradation conditions when applying it in practice.

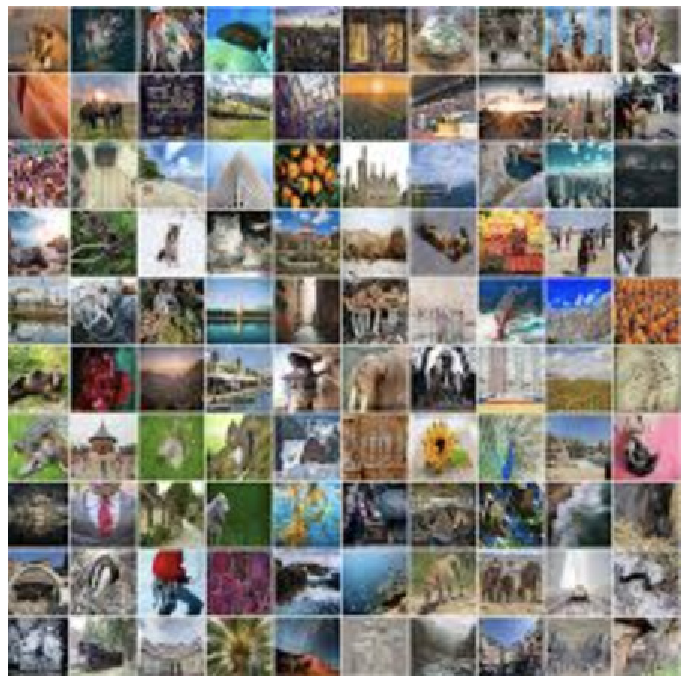

Figure 1: example of DIV2K

3 TASK THAT HAVE DONE

- By Mar. 7, Understanding the content of the paper.
- By Mar. 14, finish the investigation in literature
- By Mar. 21, Understand the code of the paper.
- By Mar. 31, reproduce the core experiments as well as the experiment partition of the work.
- By Apr. 1, submit the intermediate draft

4 TO BE DONE

- By Apr. 10, finish all the experiments and submit supplementary experiments.

- By Apr. 26, finish the final script.

REFERENCES

[1] Banham, M. R., and Katsaggelos, A. K. Digital image restoration. *IEEE signal processing magazine 14*, 2 (1997), 24–41.

[2] Bethge, J., Bartz, C., Yang, H., Chen, Y., and Meinel, C. Meliusnet: Can binary neural networks achieve mobilenet-level accuracy?, 2020.

[3] Bin Xia, Yulun Zhang, Y. W. Y. T. W. Y. R. T. L. V. G. Basic binary convolution unit for binarized image restoration network. In *conference paper* (2023).

[4] Jain, P., and Tyagi, V. Spatial and frequency domain filters for restoration of noisy images. *IETE Journal of Education 54*, 2 (2013), 108–116.

[5] Lee, C., Kim, H., Park, E., and Kim, J.-J. Insta-bnn: Binary neural network with instance-aware threshold, 2023.

[6] Liu, Z., Shen, Z., Savvides, M., and Cheng, K.-T. Reactnet: Towards precise binary neural network with generalized activation functions, 2020.

[7] Qin, H., Gong, R., Liu, X., Bai, X., Song, J., and Sebe, N. Binary neural networks: A survey. *Pattern Recognition 105* (2020), 107281.

[8] Wang, H., Wang, J., Geng, Y.-C., Song, Y., and Liu, J.-Q. Quantum image encryption based on iterative framework of frequency-spatial domain transforms. *International Journal of Theoretical Physics 56* (2017), 3029–3049.

[9] Wang, Z., Cun, X., Bao, J., Zhou, W., Liu, J., and Li, H. Uformer: A general u-shaped transformer for image restoration, 2021.

[10] Yuan, C., and Agaian, S. S. A comprehensive review of binary neural network. *Artificial Intelligence Review 56*, 11 (Mar. 2023), 12949–13013.

[11] Zhang, K., Zuo, W., Gu, S., and Zhang, L. Learning deep cnn denoiser prior for image restoration, 2017.

[12] Zong, L. *Comprehensive Study of Image Restoration Algorithms*. PhD thesis, California State University, Northridge, 2014.