Shape feature extraction for image recognition with CNN using frequency domain

Palash Dahiphale

***Abstract*— Learning shape features robustly using CNN with frequency domain deep learning for accurate image recognition.**

I. INTRODUCTION

In present days, a lot of computer vision tasks are being done by deep learning models (most popularly CNN) but, with lower computational speed and lower accuracy specially for highlighted images. Discrete Fourier transforms provide a significant speedup in the computation of convolutions in deep learning. In this work, i demonstrate that, beyond its advantages for efficient computation, the spectral domain also provides a powerful representation in which to model and train convolutional neural networks (CNNs).

II. PROCEDURE FOR REPORT SUBMISSION

A. *Why frequency domain ?*

One powerful property of frequency analysis is the operator duality between convolution in the spatial domain and element-wise multiplication in the spectral domain, which gives considerable computational ease. Moreover, the key shape information of an image is in the phase part of an image (Ref : Experiment-1) which can be extracted using frequency analysis as well as denoising of an image is an easy task in frequency domain (Ref : Experimrnt-2). The unitarity of the Fourier basis makes it convenient for the analysis of approximation loss. More specifically, Parsevals Theorem links the l2 loss between any input x and its approximation x to the corresponding loss in the frequency domain.

B. *Applications of spectral representations*

First one is spectral parametrization. In spectral parametrization, we propose the idea of learning the filters of CNNs directly in the fre- quency domain. Namely, we parametrize them as maps of complex numbers, whose discrete Fourier transforms correspond to the usual filter representations in the spatial domain. Second one is spectral pooling. Pooling refers to dimensionality reduction used in CNNs to impose a capacity bottleneck and facilitate computation. We can use the new approach to pooling called as spectral pooling. It performs dimensionality reduction by projecting onto the frequency basis set and then truncating the representation.

*This work was not supported by any organization

III. EXPERIMENTATION

I have demonstrated some experiments to support my assumptions. Further, using that assumption i have checked speed and accuracy of the particular model and done some analysis on the results. To remove the difficulties faced, again i have done some more experimentation. Using all the experiments i have done, i further checked speed and accuracy of an improved model.

- Experiment-1 To support the assumption that key information of an image lies in the phase part of an image, i have taken two images and computed FFT of both the images to separate out phase and magnitude part of it. I combined Magnitude part of image-1 with phase part of image-2 and vice versa, further i reconstructed that combined images by taking inverse FFT. Results supported our assumption.
- Experiment-2 To support the assumption that denoising can be done through frequency domain analysis which will help to separate out actual containt of an image. I have taken an image and passed through the high pass filter. Result shows outline of an image that means more information of an image in the lower frequency range of an image which supports our assumption.
- Experiment-3 In this experiment, i have computed the FFTs of an image and feed it the input of an CNN instead of an whole image. Results showed me that trained data require considerable larger storage and slower computational speed.
- Experiment-4 After doing some analysis on results of the previous experiment, i have landed on conclusion that improper pooling strategy leads to larger memory requirements. In this experiment, i have demonstrated that how use of spectral pooing avoids the problem of larger storage.
- Experiment-5 After doing some analysis on results of the experiment-3, i have landed on the conclusion that problem of slower computational speed is because of learning of filters of the CNN in spatial domain. In this experiment i have demonstrated that why to use spectral parametrization to avoid problem of slower computational speed.
- Experiment-6 After analyzing the above problems, i have used both the solution to train the model for CIFAR-10 data-set and got some interesting results i.e 81.24 percent accuracy with less no of epochs.

ACKNOWLEDGMENT

I would like to thanks Dr. Subrat Kar sir for useful discussions and assistance throughout this project. I also want to thanks Mr. Manohar kumar sir for giving me the direction and outline of the project.

References are important to the reader; therefore, each citation must be complete and correct. If at all possible, references should be commonly available publications.

REFERENCES

[1] Oppenheim's experiment section 2.6 http://cdn.intechopen.com/pdfs
[2] Phase congruency http://homepages.inf.ed.ac.uk
[3] Fast training CNN through FFTs https://arxiv.org/abs/1312.5851.
[4] Spectral representation of CNN https://arxiv.org/abs/1506.03767.
[5] Fourier CNN http://ecmlpkdd2017.ijs.si/papers/paperID11.pdf.
[6] Fast Convolutional Nets With fbfft: A GPU Performance Evaluation https://arxiv.org/abs/1412.7580.