Image De-noising using Deep Image Prior

A Synopsis Submitted
in Partial Fulfillment of the Requirements
for the Course of
Minor Project - II
in
Third-year
Bachelor of Technology
specialization
In
Artificial Intelligence and Machine Learning

Under

Mrs. Richa Choudhary
Assistant Professor - Senior Scale
Department of Informatics
UPES, Dehradun

by

Name	SAPID	Specialization
Ansh Gupta	500075765	AI & ML
Aryan Rastogi	500075858	AI & ML
Swarnim Tripathi	500075937	AI & ML

UPES
UNIVERSITY OF TOMORROW

DEPARTMENT OF INFORMATICS
SCHOOL OF COMPUTER SCIENCE
UNIVERSITY OF PETROLEUM AND ENERGY STUDIES, BIDHOLI,
DEHRADUN, UTTRAKHAND, INDIA

1. Introduction

According to InfoTrends, 1.2 trillion digital photos will be taken worldwide this year, which equates to roughly 160 photos for every one of the earth's 7.5 billion people. Thereby, image degradation is a major issue in the digital world where images are being constantly captured, shared, and stored. Image degradation can be defined as the loss of image quality owing to any circumstance that causes blurring or a drop in image quality. Degradation comes in many forms such as motion blur, noise, and camera misfocus. In cases like motion blur, it is possible to come up with a very good estimate of the actual blurring function and "undo" the blur to restore the original image. Degradation can happen at any stage of storing, transmitting, or capturing, and clearly restoring such images is critical.

As the models in the paper “Deep Image Prior”[1] stand, the restored images or the high-resolution images do not surpass the state-of-the-art techniques in denoising like CBM3D[2] or SRRESnet[3]. However, the Deep Image Prior technique finds itself very handy when trained architectures cannot be used. A specific instance can be images of some novel virus whose predominant images are not available. The Deep Image Prior models can be tweaked for the specific image to get useful higher resolution images. However, it is true that doing this in its own right is not easy, and understanding the baseline convolutional neural network is tough.

We will be emphasizing on Deep Image Prior, which is a type of Convolutional Neural Network which is used to improve an image with no prior training data other than the image itself. A neural network is randomly seeded and then used prior to solving inverse problems like noise reduction, super-resolution, and inpainting. The structure of a convolutional image generator, rather than any previously learned capabilities, captures image statistics.

A few interesting things that came up in the paper were the consideration of how the hand-crafted network behaved as the prior for the image. Whatever information is stored about the task or the image is in the network and thus called a prior. The degraded image or the image to be super magnified features only in the error function which is used for gradient descent by the network provided the network has not been fine-tuned for the image. The second thing was if gradient descent was run for long enough they converged to the respective input images so it was imperative to identify the right stopping point. To summarize, the paper indicates that hand-crafted network architectures can be adequate to solve image restoration tasks up to certain levels of accuracy (without any explicit training).

This project has largely been motivated by the desire of restoring degraded images. In this reference, we stumbled across the paper “Deep Image Prior”[1] that promises to carry out image processing tasks without any training data set. This was done by handcrafting a neural network for the specific task and the specific image. In this project, we have tried to understand the making of such architecture for two image processing tasks: Denoising and Super-Resolution.

2. Motivation

Image denoising is a basic difficulty in the field of image processing and computer vision, where the underlying objective is to estimate the original image by suppressing noise from a noise-contaminated version of the image. Image noise may be created by a variety of intrinsic(i.e, sensor) and extrinsic(i.e, environmental) factors that are frequently unavoidable in real-world scenarios. While various algorithms have been developed for image denoising, the topic of image noise suppression remains an outstanding challenge, particularly in scenarios where the images are acquired under poor conditions where the noise level is very high. One of the most significant motivations for this project is to denoise satellite images. Images taken by cameras setup in satellites are sometimes noisy, and due to absence of any training data, those images cannot be denoised. By using the deep image prior technique, we can tackle this issue as this technique does not need any training data.

3. Related Work

The techniques that we are using are related to image restoration and synthesis methods based on learnable ConvNets and deep image prior. Here, we review other lines of work related to our approach.

The zero-shot super-resolution technique [6] is an interesting parallel effort with clear ties to our approach in this group, which trains a feed-forward super-resolution ConvNet based on a synthetic dataset constructed from the patches of a single picture. While obviously related, the technique [6] is slightly complimentary in that it uses self-similarities across several scales of the same picture, whereas our approach uses self-similarities within the same scale (at multiple scales).

The work [7] develops a latent variable model, which is parametrized using a convolutional neural network. The method [8] attempts to reconstruct an image from an event-based camera and using a deep image prior framework to estimate the sensor's ego-motion. The approach [9] successfully employs deep image prior to defending from adversarial assaults. In [10] deep image prior is also employed to accomplish phase retrieval for Fourierptychography.

In [11] study, they explore the key difficulties surrounding Convolutional Neural Networks (CNNs) and explain how each parameter affects network performance. The convolution layer is the most crucial layer in CNN, as it takes up the majority of the network's time. The number of levels in a network has an impact on its performance. However, as the number of layers grows, so does the amount of time it takes to train and test the network. Today, CNN is regarded as a powerful machine learning technique for a variety of applications, including face detection and image recognition, video recognition, and speech recognition.

There are a couple of popular techniques is image restoration and image denoising like • LapSRN [4] (Trained) • SRRESnet- (Trained) and • CMB3D • NLM [5] respectively.

4. Proposed Method

In the proposed approach, a Convolutional Neural Network will be used for the purpose of denoising the image. In the whole method, we will be considering three images,

a) Ground Truth: Image with no noise
b) Noisy Image: Image with some noise
c) Image with pure noise: An image with 100 percent noise

Goal: To generate an output image which is near to the ground truth(a).

Technique: For training the machine learning model, we will use the noisy image(b). For the purpose of denoising, we will be making use of the image with pure noise(c). We will try to run the model for some iterations and generate our output image.

Finally, for model evaluation, we compare our generated output image with the ground truth(a) and determine the loss that defines the accuracy of the built model.

5. Methodology

Datasets used: The Deep Image Prior model is not trained on any data set, hence there is no training data set used in the model. For model building, we will run the code on some random images from the internet.

Procedure: Let x be the clean image, x_0 be the degraded image, and x^* be the restored image. Consider,

$$x = f_\theta(z)$$

From our prior knowledge of Machine Learning, we know to get optimum x following equation need to be satisfied

$$x^* = arg\ maxp(x_0 \mid x)$$

using the Bayesian rule,

$$p(x|x_0) = [p(x_0|x).p(x)] \,/\, p(x_0)$$

Hence,

$$\begin{aligned} x^* &= arg\ maxp(\mathrm{x} \mid x) \\ &= arg\ min - log\ p(x_0 \mid x) - log\ p(x) \\ &= arg\ min\ E(x;x_0) + R(x) \\ x^* &= min_x\ E(x;x_0) + R(x) \end{aligned}$$

$R(x)$ is a regularizer. Now on taking $R(x) = 0$ we get,

$$\theta^* = arg_\theta\ minE(f_\theta(z);\ x_0), \quad x^* = f_{\theta^*}(z)$$

Our aim will be to achieve the above equation. The optimal theta, θ^* is obtained by using an optimizer such as gradient descent, starting from a randomly initialized x with random noise and theta.

A. Super-Resolution

In super resolution, the input is a low resolution image x_0, and a magnifying factor t such that the output is a high resolution image x.

We will talk about how super-resolution will be implemented in the project. As pre-processing steps, the original image will be trimmed at the borders so that both the height and width of the image is divisible by 32. (This is because of 5 downsampling layers in the magnifying neural net which reduces height, width by a factor of 2) Note that a Low Resolution image is not taken as input but the image with the original specifications is reduced by the factor of desired magnification and this resized image is then used to generate the desired high resolution image.

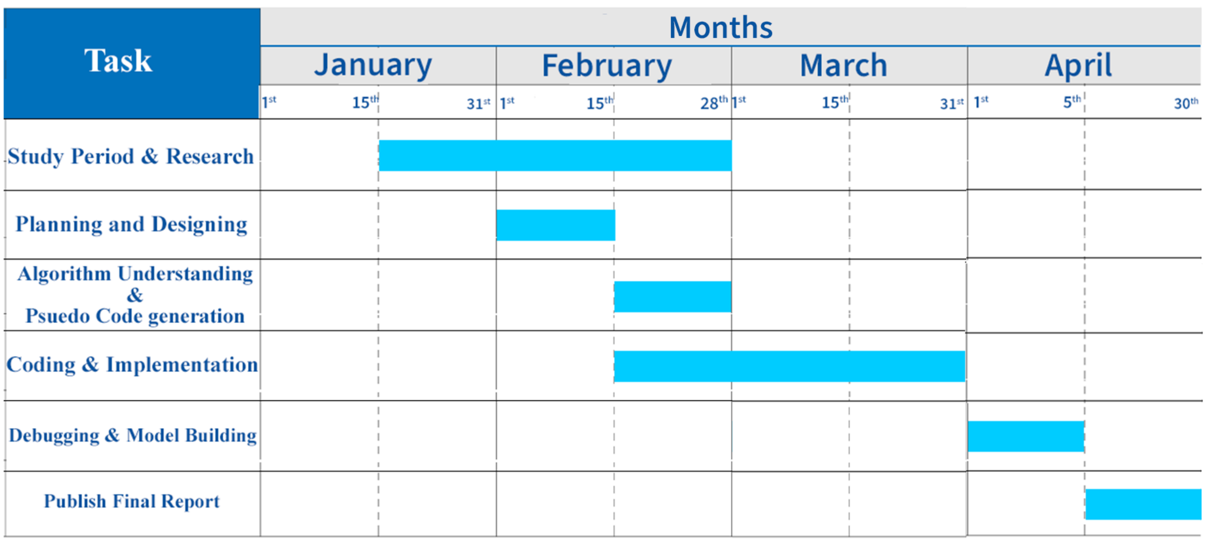
B. Denoising

Denoising the process of obtaining a clear image, x from a noisy image, x_0. Consider a model, x0 = x+ where, distribution is the noise and may optionally be incorporated into the model, though this process works well in blind denoising.

We will take some random noise in the form of an image (z) and feed that into the model. The model returns us some output that is the map of the input (z). The output is $f_\theta(z)$ for some set of params that we call θ. We optimize over this set of parameters (θ) such that the map $f_\theta(z)$ more closely resembles the noisy image, i.e, the entropy loss between the output of the model and the input noisy image is minimized.

6. Plan of work

Gantt Chart



References

[1] Dmitry Ulyanov, Andrea Vedaldi,and Victor Lempitsky, Deep Image Prior, 2017.

[2] Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Transactions on image processing 16(8), 2080-2095 (2007)

[3] Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J.,Wang, Z., Shi, W.: Photo-realistic single image super-resolution using a generative adversarial network. In:CVPR. IEEE Computer Society (2017)

[4] Lai, W.S., Huang, J.B., Ahuja, N., Yang, M.H.: Deep laplacian pyramid networks for fast and accurate super-resolution. In: CVPR. IEEE Computer Society (2017)

[5] Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: Proc. CVP vol. 2, pp. 60-65.IEEE Computer Society (2005)

[6]Assaf Shocher Nadav Cohen, M.I.: "zero-shot" super- resolution using deep internal learning. In: CVPR. IEEE Computer Society (2018)

[7]Athar, S., Burnaev, E., Lempitsky, V.S.: Latent convolu- tional models. CoRR (2018)

[8] Shedligeri, P.A., Shah, K., Kumar, D., Mitra, K.: Photo-realistic image reconstruction from hybrid intensity and event based sensor. CoRR (2018)

[9] Ilyas, A., Jalal, A., Asteri, E., Daskalakis, C., Dimakis, A.G.: The robust manifold defense: Adversarial training using generative models. CoRR (2017)

[10] Boominathan, L., Maniparambil, M., Gupta, H., Babu- rajan, R., Mitra, K.: Phase retrieval for fourier ptychog- raphy under varying amount of measurements. CoRR (2018)

[11] S. Albawi, T. A. Mohammed and S. Al-Zawi, :Understanding of a convolutional neural network, ICET (2017)