Restoration of Compressed Images by Deep Image Prior

Pawar Prathamesh, Wadhwa Eshan
Northeastern University, Boston
pawar.prath@northeastern.edu, wadhwa.e@northeastern.edu

Abstract

Deep Convolutional Networks have been widely utilized as a fundamental tool for image generation, classification, and segmentation. In this report, we explore the application of the techniques proposed in the paper Deep Image Prior [1] to investigate the use of CNN architectures in low-level image transformation tasks, such as restoration and denoising. Specifically, we focus on the application of image restoration following lossy compression, replicating and analyzing the methods described in the paper.

1 Introduction

In the paper Deep Image Prior [1], the authors demonstrate that Convolutional Neural Networks (CNNs), while powerful due to the prior knowledge learned from training on large datasets, can also be effectively utilized without leveraging image priors. In this contrast, CNNs can capture significant learning solely based on their architecture. To illustrate this, the authors used an untrained ConvNet, fitting it to a single image. This approach involves using a neural network to restore damaged or low-quality images, but unlike traditional methods where the network is trained on large datasets, this method works with one degraded image at a time. The weights of the network are initially set randomly and then iteratively adjusted to match the specific input image. The network's structure is designed by humans to fit the task at hand, and the restoration relies solely on the information from the single input image and the network's pre-designed structure. In this report, we apply the technique described in this paper to restore an image degraded by lossy compression, with the goal of bringing it as close to the original as possible.

2 Prior Work

The authors of *Deep Image Prior* theorize that the parameters of the model are not learned from the prior distribution of images $p(x)$ in the training dataset. Instead, a significant amount of learning occurs due to the structure of the network itself. The authors interpret the network's parameters as a parametrization of the image x based on a code vector z, where the relationship is defined as:

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

Here, x represents the image, z is the code vector, and θ denotes the parameters of the network.

The paper demonstrates that even without traditional training, the network inherently captures low-level statistics of natural images. This capability arises from the local, translation-invariant nature of convolutions and the hierarchical structure of multi-scale operations. These characteristics enable the network to model conditional image distributions $p(x \mid x_0)$, which are essential for solving inverse problems such as denoising, super-resolution, and inpainting.

Traditional image restoration methods typically minimize an energy function:

$$x^* = \arg\min_x E(x; x_0) + R(x)$$

In this formula, $E(x; x_0)$ represents a task-specific data term, and $R(x)$ is an explicit regularizer that captures natural image regularities (e.g., the Total Variation (TV) norm). In contrast, this work introduces an innovative approach by eliminating the explicit regularizer $R(x)$, relying instead on the implicit prior encoded in the network's architecture. The restoration process involves optimizing the network parameters θ, starting from a random initialization, in order to minimize:

$$E(f_\theta(z); x_0)$$

Here, $f_\theta(z)$ generates the restored image x^*, which is the outcome of the optimization procedure. This approach allows the network to directly capture and restore image details without the need for predefined regularizers, relying solely on the inherent properties of the network and its structure.

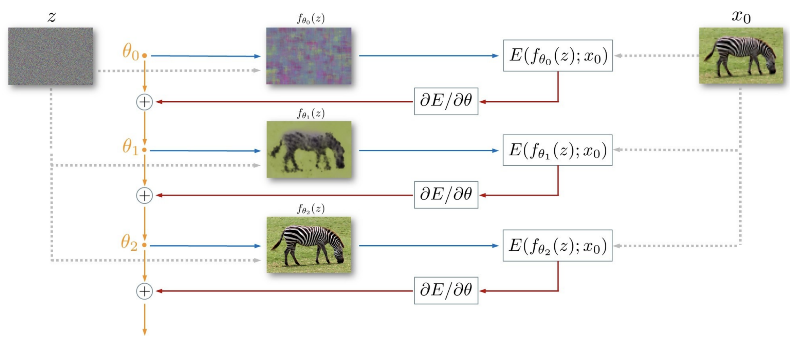

Figure 1: Image restoration using the deep image prior

2.1 Image Space Visualization

In Fig. 2, we can observe a visualization of the Deep Image Space. Here, the authors define a ground truth image x_{gt}, which is degraded to x_0 as the target image. It is evident that the ground truth x_{gt} has a non-zero cost, i.e., $E(x_{gt}, x_0) > 0$.

When using a deep image prior, extended optimization may lead to a solution with nearly zero cost, which can deviate significantly from the ground truth x_{gt}. However, during the optimization process, the path comes close to x_{gt} at one point. Using early stopping, we stop the optimization at a specific time t_3, we can obtain a good solution.

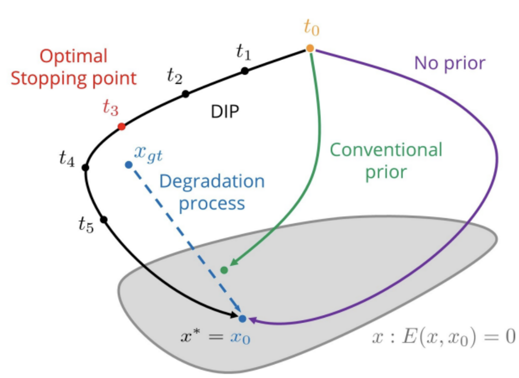

Figure 2: Image Space Visualization

2.2 Neural Network Structure

One of the most significant contributions of the authors is showing that the untrained network's architecture, particularly a U-Net-like hourglass with skip connections, provides a strong prior. This implicit prior is shown to perform competitively with, and sometimes surpass, traditional priors and approaches trained on data. A "Hourglass" shaped UNet with skip connection has the ability to combine structures at multiple scales as it happens in natural images. It is also helps model recover spatial information that may have been lost in down-sampling[2]. The network's millions of parameters and randomly initialized latent tensor z make this method broadly applicable across diverse image restoration tasks.

3 Method

3.1 Model Architecture

Using the model recommended by the authors, we designed a CNN with UNet architecture and skip connections at every "block". By using skip connections, we can preserve the spatial information much more effectively than a normal CNN.

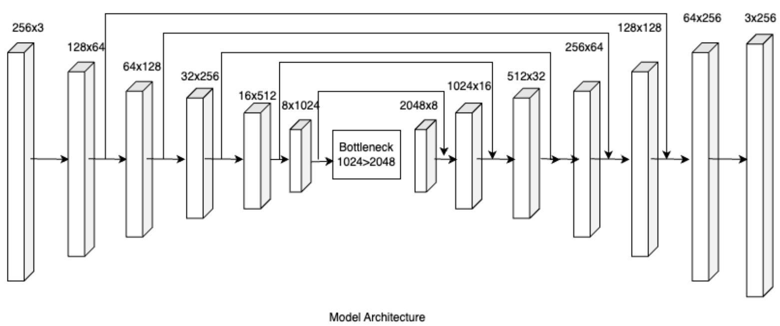

Figure 3: Model Architecture for UNet

3.2 Experiments

We use this model to restore a compressed image to its original image[3]. In order to test the performance, we first compress an image to 25% of its original size. We then generate random noise and use that as an input to the model. The degraded (compressed) image is used as a target image for the network. We optimize the model for this setup and record the PSNR values of the output of the model to both the original and degraded image. However we only use the degraded image for optimization.

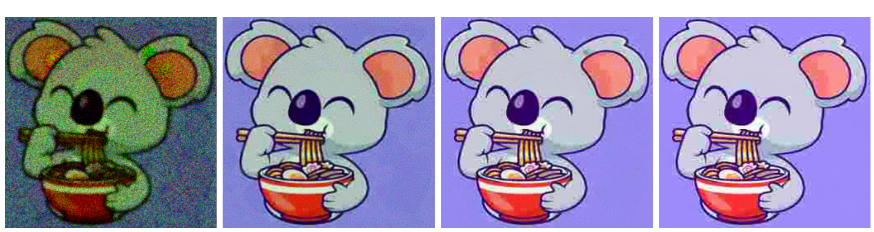

Figure 4: Iterations: (a) 1000 (b) 10000 (c) 15000 (d) 20000

Figure 5: Iterations: (e) 30000 (f) 35000 (g) 40000 (h) compressed

4 Results

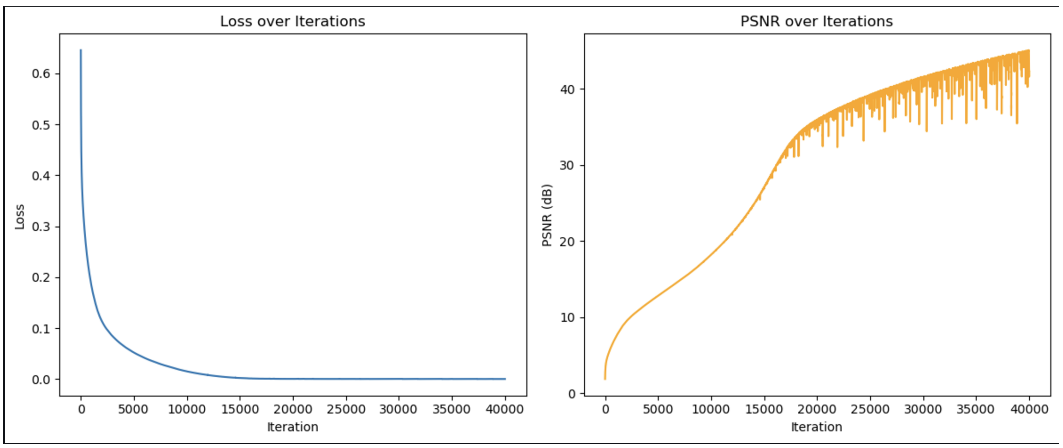

Figure 6: Training Loss and PSNR trend

We find that the output of the model is closest to the original image at around 35000 Iterations.

5 Conclusions

After running extensive experiments we can conclude that, an untrained CNN can be used to restore a compressed image to its original form. We concur with the authors' claim that fitting a randomly initialized ConvNet can be used as a common solution to a restoration application. We found evidence that UNet architecture with skip connections tend to impose self-similarity on multiple scales making them suitable for restoration applications.

References

1. Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2020). Deep image prior. International Journal of Computer Vision.

2. Adaloglou, N. (2020). Intuitive explanation of skip connections in deep learning. The AI Summer. Retrieved December 3, 2024, from https://theaisummer.com/skip-connections/

3. Image Source: freepik.com