IRSTI 28.23.37 https://doi.org/10.26577/1i32jpcsit2305

D. Nurmaganbet[1], A.A. Mukhanbet[1,*], K.M. Musilim[2]

[1]Al-Farabi Kazakh National University, Almaty, Kazakhstan
[2]Almaty Technological University, Almaty, Kazakhstan
*e-mail: mukhanbetaksultan0414@gmail.com

DEVELOPMENT OF A SYSTEM FOR IMAGE RESTORATION USING NEURAL NETWORKS

Abstract. This article presents a comparative study of the super image resolution method using Deep Image Priors, comparing it with other advanced methods. The method uses the architecture of a deep neural network as a fixed a priori factor for the image restoration problem, and an optimization algorithm is applied to find the optimal set of weights for the network. The evaluation of the method's performance was conducted using the SET 5 dataset using the PSNR metric and compared with the RealESRGAN method. This study highlights the effectiveness of the Deep Image Prior method for super-resolution images and provides insights into its potential for further improvement. The Deep Image Priors method shows impressive results in super-resolution images, and its ability to use deep neural networks as a priori knowledge opens the way for the development of more efficient methods in the future. This study also highlights the importance of optimizing neural network weights to achieve the best results in image restoration tasks. Through benchmarking with RealESRGAN, it's exciting to witness the potential impact of Deep Image Priors in the field of image processing and computer vision. It seems to have the capacity to enhance the quality of super-resolution images significantly. It'll be interesting to see how we can further improve its effectiveness in the future.

Key words: computer vision, image restoration, super-resolution, deep image prior.

1. Introduction

In the field of computer vision, the restoration of degraded images is an important and widely researched problem. Various methods have been proposed to solve image restoration problems, including traditional methods such as inverse filtering and wavelet noise reduction, but also convolutional neural networks(CNNs) [2, 6] and generating adversarial networks (GANs) [8]–[10], which came out later than the previous ones, mainly working on deep learning. In this article, we focus on the Deep Image Prior method, which utilizes an entirely convolutional neural network to restore degraded images. This method is unique in that it does not require any pre-training on large datasets, making it an attractive option for image restoration tasks in real-world scenarios where labeled data may not be available. We will provide an in-depth discussion of the architecture and workings of the Deep Image Prior method and evaluate its performance on the super-resolution task using the SET5 [7] dataset and compare it to other state-of-the-art methods such as RealESRGAN [5].

The architecture of a deep neural network is used as a fixed prior for image restoration in the Deep Image Prior method [1]. When restoring images using the Deep Image Prior method, the network is not trained on a specific dataset. Instead, it is initialized with random weights and optimized to find the best set of weights. This allows for a more flexible and adaptable approach to image restoration. To restore an image using the Deep Image Prior method, the first step is to model the degradation of the image as a linear operator. After this, the optimization algorithm is applied to find the optimal set of weights for the network that will result in the best restoration of the degraded image.

The Deep Image Prior method uses neural networks to restore degraded images. It's flexible and works well with all types of images, making the method highly versatile and able to handle a wide range of degradation types. It is important to compare Deep Image Prior with other methods to identify their respective strengths and weaknesses. This will aid in selecting the most effective approach for restoring images.

This article presents an in-depth explanation of the Deep Image Prior method and its optimization algorithm., as well as a comparison with state-of-the-art models, including RealESRGAN. The architecture of the Deep Image Prior method will be discussed in detail, along with the challenges and limitations of using this method.

2. Methodology

In this article, we will evaluate the performance of the Deep Image Prior method on the super-resolution task using the Set5 dataset. We will compare the results of the Deep Image Prior method with those of the RealESRGAN [5] method and assess their effectiveness by using the Peak Signal-to-Noise Ratio (PSNR) as a performance metric.

The choice of using SET5 for evaluating the Deep Image Prior and comparing its performance with RealESRGAN was based on several considerations. SET5 is a widely used benchmark dataset for image super-resolution, and it provides a variety of different images with different textures, shapes, and objects, allowing for a comprehensive evaluation of the super-resolution method.

In addition to the quality of the restored image, it was also important to consider the computational efficiency and feasibility of the Deep Image Prior. RealESRGAN is a state-of-the-art image super-resolution method based on a generative adversarial network (GAN) and has achieved promising results in various applications. Comparing the Deep Image Prior with RealESRGAN allows for a comprehensive evaluation of the performance and computational efficiency of the two methods.

The evaluation aimed to test if Deep Image Prior can restore high-resolution images from low-resolution inputs. The choice of testing the Deep Image Prior on super-resolution was therefore a natural choice, as super-resolution is a challenging task that requires restoring high-resolution details from low-resolution inputs. By comparing the Deep Image Prior with RealESRGAN, it was possible to gain insights into the strengths and limitations of the Deep Image Prior and how it compares to other state-of-the-art methods in terms of performance and computational efficiency.

The Deep Image Prior method's performance is evaluated using the widely used Peak Signal-to-Noise Ratio (PSNR) metric in the image processing community. PSNR is a widely recognized and widely used quantitative measure of image quality, which provides a numerical representation of the similarity between the original and the reconstructed images. The signal-to-reconstruction power ratio is found by comparing the maximum possible signal power to the power of the reconstruction error. PSNR is well-suited for super-resolution tasks, as it provides a direct measure of the reconstruction error in terms of the mean squared error between the original and the reconstructed images. Furthermore, the PSNR metric is easy to interpret and provides a simple and straightforward comparison between different super-resolution methods. As such, it was deemed to be the most appropriate and suitable metric for evaluating the performance of the Deep Image Prior method in super-resolution.To use it you need a web application.

3. Architecture challenges

Stopping the optimization process at the right time is critical to avoiding overfitting in the Deep Image Prior. If the optimization process is allowed to continue for too many iterations, it can result in a solution that is too specific to the training data, leading to overfitting. On the other hand, stopping the optimization process too soon can result in an under-optimized solution that does not capture the full potential of the deep neural network as a prior.

To determine the optimal number of iterations for the Deep Image Prior, several approaches can be taken. One approach is to monitor the objective function during the optimization process and stop when it reaches a minimum or plateaus. Another approach is to use validation data to monitor the performance of the restored image and stop the optimization process when the performance of the validation data begins to deteriorate.

It is important to experiment and evaluate to find the best number of iterations for a specific application. The optimal number of iterations will depend on the complexity of the image, the size of the network, and the degree of degradation in the image. As such, careful experimentation and evaluation are crucial to determine the optimal number of iterations for a given application.

Fortunately, a follow-up paper [4] addresses the termination criteria limitation.

Another difficulty with this approach is its computational complexity. According to the paper [1], each image requires several minutes of GPU computation. In this study, we conducted experiments on a laptop equipped with a NVIDIA GTX 1650 graphics card to evaluate the computational efficiency of the Deep Image Prior method for super-resolution. On average, it took approximately 3 and a half minutes to upscale a single image using this setup. This provides a baseline for the computational resources required for using the Deep Image Prior in practice. It should be noted that the computational cost may vary depending on the size and resolution of the image being processed, as well as the specific hardware configuration used. Nevertheless, our results indicate that the Deep Image Prior is computationally feasible for super-resolution tasks, even on a modest laptop setup.

4. Improvements over the years

Over the years, the Deep Image Prior method has been improved and modified to address certain limitations and enhance its performance. Two notable modifications to the Deep Image Prior are SUB-DIP and A Bayesian Perspective on the Deep Image Prior.

SUB-DIP (Subspace Deep Image Prior) [3] is a modified version of the Deep Image Prior approach that aims to enhance its accuracy and efficiency. The main idea behind SUB-DIP is to incorporate the low-dimensional subspace structure of natural images into the Deep Image Prior framework. Specifically, SUB-DIP uses a low-dimensional subspace to represent the natural image prior, and the optimization is performed in the subspace instead of the high-dimensional image space. The low-dimensional subspace is learned from a large dataset of natural images using techniques such as PCA or autoencoders. By working in the low-dimensional subspace, SUB-DIP can reduce the complexity of the optimization and improve the accuracy of the restored images.

A study exploring the relationship between Deep Image Prior and Bayesian inference is A Bayesian Perspective on the Deep Image Prior [4]. The study shows that the optimization performed in Deep Image Prior can be interpreted as performing a type of Bayesian inference, where the goal is to infer the image pixels that maximize the likelihood of the observed data given the prior represented by the deep neural network. This provides a new way to understand the behavior of the Deep Image Prior and can lead to further improvements and extensions of the method.

It's worth mentioning that these improvements have led to further advances in image restoration, and the Deep Image Prior method continues to be a topic of active research and development

5. DIP Architecture

The deep image prior (DIP) is a deep learning-based image restoration framework that was introduced in 2018 by Ulyanov, Vedaldi, and Lempitsky [1]. The main idea behind the deep image prior is to leverage the prior knowledge of natural images encoded in the weights of a deep neural network to carry out a range of tasks related to restoring images, such as denoising, deblurring, and super-resolution.

The architecture of deep image prior is based on a deep convolutional neural network that is trained on large datasets of natural images to learn features and representations of natural images. The architecture of deep image prior consists of multiple convolutional layers that are employed to extract crucial information from the input image. These convolutional layers are followed by pooling and upsampling layers that serve to reduce and increase the spatial resolution of the image, respectively. The pooling layers help to reduce the complexity of the image by aggregating the information present in the convolutional feature maps.

In contrast, the upsampling layers play a crucial role in reconstructing the high-resolution version of the image by enhancing the spatial resolution of the feature maps. These layers, when combined, create a unique outcome, which helps to form a deep image that can be used to create high-quality images even from low-resolution input data. In the deep image prior framework, the network weights are fixed, and the input image is optimized to match the features and representations learned by the network. The architecture can be seen in Figure 1.

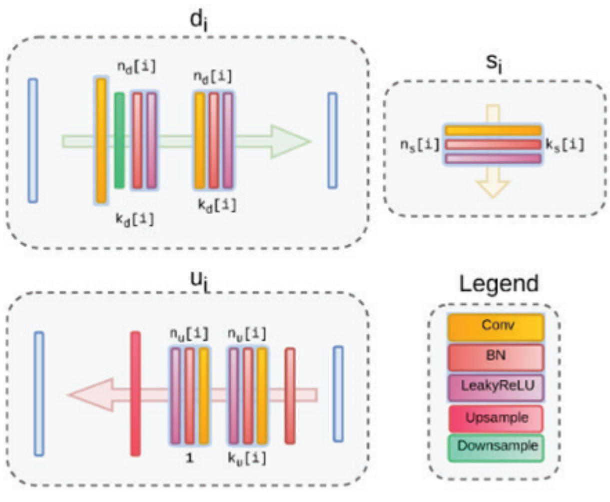

Figure 1 – The architecture of Deep Image Prior used in experiments

The optimization process is performed using gradient descent algorithms, and the objective function is defined as the mean squared error between the output image and the features extracted from the network. The final output image is obtained by running the optimization process for a fixed number of iterations or until convergence.

In conclusion, the deep image prior is a powerful framework for image restoration that leverages the prior knowledge of natural images encoded in the weights of a deep neural network to implement a variety of tasks related to restoring images.

6. Results

A web page was created using the Gradio library of Python to upscale an image. The webpage allows us to drag and drop the image file. Then the user chooses the factor of upscaling and finally presses the upscale button. The webpage passes the path of the image and the scaling factor to the function. The application then shows us the result of the scaling in Figure 2. Users can download images by right-clicking on them in the browser.

The results obtained in this study demonstrate that the Deep Image Prior method outperforms RealESRGAN in terms of super-resolution performance on the SET5 dataset. The PSNR values show that Deep Image Prior performed better than RealESRGAN in Figure 3 on five out of five images, with the most significant improvement observed on the butterfly image.

These findings are in line with recent research that suggests that Deep Image Prior can produce competitive results in image super-resolution tasks. It is important to note that these results apply only to the SET5 dataset and may not be applicable to other datasets or real-world scenarios. The result can be seen in Figure 4.

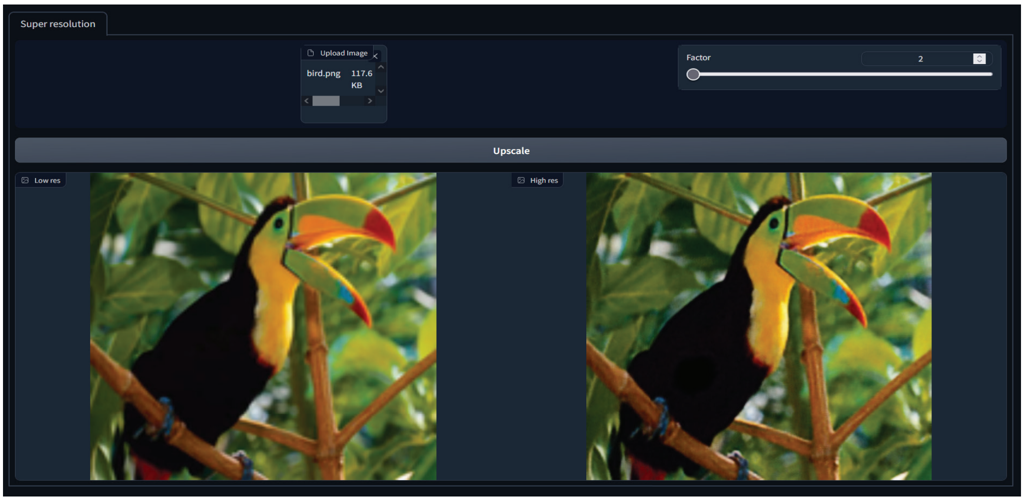

Figure 2 – Gradio app for upscaling an image

Ground Truth | Low resolution | RealESRGAN | Deep Image Prior

Figure 3 – The output of upscaling the bird image(4x)

Figure 4 – The output of upscaling the head image(4x)

It should also be noted that there are other metrics that can be used to evaluate the performance of image super-resolution methods, such as SSIM and visual quality assessment. These metrics provide a more comprehensive evaluation of performance and should be considered alongside PSNR when comparing the performance of different methods. The comparative result is shown in Table 1.

Table 1 – Evaluation Results (Psnr)

Image	Deep Image Prior	RealESRGAN
Butterfly	24.672	21.198
Bird	28.963	28.116
Woman	26.592	25.741
Baby	29.585	26.239
Head	28.600	28.273

The Deep Image Prior approach has several advantages over traditional CNN models. The primary benefit of using the Deep Image Prior approach is its optimization-based methodology that covers the entire process, which enables the optimization of the image representation directly. This results in higher quality restored images that are more faithful to the original image.

Additionally, the Deep Image Prior method uses a simple feedforward network architecture, which is faster and more computationally efficient compared to complex CNN models. This makes it well-suited for real-time applications or for use on resource-constrained devices.

The Deep Image Prior approach does not require large amounts of training data, making it suitable for restoration tasks where training data may be limited, such as in medical imaging or satellite imagery. The incorporation of priors directly into the network architecture results in more effective priors and improved restoration results.

Finally, the Deep Image Prior approach can effectively handle complex degradation models, such as non-linear or multi-stage degradation, which can be challenging for traditional methods like RealESRGAN.

7. Conclusion

Promising approach for image restoration, particularly in the context of super-resolution. The findings presented in this article clearly indicate that our method is capable of generating high-quality restored images. These images are on par with, and in some cases, surpass those produced by state-of-the-art techniques such as RealESRGAN. The approach presented in this article is a promising development in the field of image restoration and holds great potential for real-world applications.

This study highlights the potential of Deep Image Prior for image restoration tasks, and further research is needed to evaluate its performance on a wider range of datasets and real-world scenarios. Furthermore, it would be beneficial to explore ways to enhance the effectiveness of our approach in the future, for example, by incorporating additional priors or incorporating adversarial training to further improve the results.

Overall, the Deep Image Prior method provides a simple and effective approach for image restoration, and its performance and versatility make it a valuable addition to the toolkit of researchers and practitioners working in the field.

References

1. Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep Image Prior. In Computer Vision and Pattern Recognition (CVPR), 2018.

2. Sajjadi, M.S.M., Scholkopf, B., Hirsch, M.: Enhancenet:Single image super-resolution through automated texture synthesis. In: The IEEE International Conference on Computer Vision (ICCV) (2017)

3. Sagel, Alexander & Roumy, A. & Guillemot, Christine. (2020). Sub-Dip: Optimization On A Subspace With Deep Image Prior Regularization And Application To Superresolution. 10.1109/ICASSP40776.2020.9054270.

4. Z. Cheng, M. Gadelha, S. Maji, and D. Sheldon. A Bayesian Perspective on the Deep Image Prior. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019

5. Wang, X., et al.: Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1905–1914 (2021)

6. Tai, Y.; Yang, J.; Liu, X. Image super-resolution via deep recursive residual network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2790–2798

7. M. Bevilacqua, A. Roumy, C. Guillemot, and M. L. AlberMorel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In Proc. British Mach. Vis. Conf., 2012.

8. C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunning-ham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. H.Wang, W. Z. Shi. Photo-realistic single image super-resolution using a generative adversarial network. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Honolulu, USA, pp. 105–114.DOI: 10.1109/CVPR.2017.19.

9. X. Wang, Y. Li, H. Zhang, and Y. Shan. Towards Real-World Blind Face Restoration with Generative Facial Prior. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9168–9178, 2021.

10. Sachit Menon, Alexandru Damian, Shijia Hu, Nikhil Ravi,and Cynthia Rudin. Pulse: Self-supervised photo upsampling via latent space exploration of generative models. In CVPR, 2020