**Report**

**Abstract**

This study investigates the classical problem of image super-resolution by comparing three distinct methods: Deep Image Prior (DIP), Generative Adversarial Networks (GAN), and Implicit Neural Representations (SIREN). Using the DIV2K dataset, each method’s performance was evaluated under standard conditions, with added noise degradation and further downscaling for comparison. Evaluation metrics included PSNR and SSIM, which quantify image reconstruction quality. Visual analysis and performance metrics were also assessed for the SIREN model.

**1. Introduction**

Image super-resolution is a computer vision technique aimed at enhancing the resolution of low-resolution images, with applications in fields like medical imaging and video processing. This report examines three methods—Deep Image Prior, Generative Adversarial Networks (GAN), and Implicit Neural Representations (SIREN)—to assess their effectiveness in reconstructing high-resolution images, managing noise degradation, and handling further downscaling.

**2. Selected Methods**

The three selected methods each represent different algorithmic frameworks: model-based methods (DIP), generative models (GAN), and implicit neural representations (SIREN). These methods were chosen to observe the performance differences across diverse approaches to super-resolution.

**2.1 Method 1: Deep Image Prior (DIP)**

Deep Image Prior (DIP) is a model-based approach that leverages the architecture of convolutional neural networks (CNNs) to learn the structure of a specific image without requiring any training data. By exploiting the inherent structural properties of CNNs, DIP can recover high-frequency details in tasks such as denoising and super-resolution, even when data is scarce.

**2.2 Method 2: Generative Adversarial Networks (GANs)**

Generative Adversarial Networks (GANs) have gained considerable attention in image generation tasks. In this experiment, a GAN framework consisting of a generator and a discriminator was used. The generator aims to produce high-resolution images from noise, while the discriminator evaluates their authenticity. By training the generator to deceive the discriminator, GANs are capable of producing high-quality images. Here, GANs were employed to handle denoising, with performance evaluated under various noise levels and downscaling.

**2.3 Optional Method 3: Implicit Neural Representations (SIREN)**

The SIREN model, which uses sinusoidal activation functions to implement implicit neural representations, can effectively fit continuous functions (like images). It is particularly adept at representing high-frequency details, enabling it to represent images in a continuous space rather than relying on an explicit pixel grid. SIREN shows promise in reconstructing fine details in image restoration tasks.

**3. Experimental Design**

**Dataset and Evaluation Metrics**

We used the widely-adopted DIV2K dataset to perform image super-resolution experiments. Evaluation metrics included:

* **PSNR (Peak Signal-to-Noise Ratio):** Measures the quality of reconstructed images by comparing them to the original high-resolution images.
* **SSIM (Structural Similarity Index):** Assesses image quality degradation by examining changes in structural information, brightness, and contrast.

**Training**

Each method was trained on a subset of the DIV2K training images. In Task 2.2, we added noise degradation with varying noise levels (sigma=10, 15, and 20). Task 2.3 involved further downscaling the x8 images to evaluate the models’ robustness.

**Validation**

Each method was validated on a subset of the DIV2K validation set, evaluating performance on unseen images. The performance of DIP and GAN was measured using PSNR and SSIM, while SIREN was assessed through visual analysis of reconstructed images.

**Qualitative and Quantitative Comparison**

**Deep Image Prior (DIP)**

* **Standard Validation**: PSNR: 6.74, SSIM: 0.017.
* **Task 2.2 (Noise Degradation)**:
  + Sigma=10: PSNR: 8.27, SSIM: 0.187.
  + Sigma=15: PSNR: 9.15, SSIM: 0.247.
  + Sigma=20: PSNR: 6.74, SSIM: 0.017.
* **Task 2.3 (Downscaling to x16)**: PSNR: 7.43, SSIM: 0.097.

**Generative Adversarial Networks (GANs)**

* **Standard Validation (Epoch 10)**:
  + Noisy vs. Original: MSE: 0.0624, PSNR: 12.05.
  + Denoised vs. Original: MSE: 0.1329, PSNR: 8.76.
* **Task 2.2 (Noise Degradation)**:
  + Sigma=10: Noisy vs. Original: MSE: 0.3126, PSNR: 5.05; Denoised vs. Original: MSE: 0.1503, PSNR: 8.23.
  + Sigma=15: Noisy vs. Original: MSE: 0.3164, PSNR: 5.00; Denoised vs. Original: MSE: 0.1523, PSNR: 8.17.
  + Sigma=20: Noisy vs. Original: MSE: 0.3188, PSNR: 4.96; Denoised vs. Original: MSE: 0.1516, PSNR: 8.19.
* **Task 2.3 (Further Downscaling to x16)**:
  + Downscaled vs. Original: MSE: 0.0026, PSNR: 25.90.
  + Denoised vs. Original: MSE: 0.1324, PSNR: 8.78.

**Implicit Neural Representations (SIREN)**

SIREN’s performance on DIV2K images was evaluated qualitatively. SIREN effectively captured high-frequency details, yielding visually clear reconstructions. However, its quantitative performance in noise handling requires further exploration, as it lacks traditional metrics like DIP and GANs.

|  |  |  |
| --- | --- | --- |
|  |  |  |

**4. Conclusions**

This study highlights the strengths and limitations of the three methods in the context of image super-resolution:

* **Deep Image Prior (DIP)** performs well on noiseless images but struggles with high noise and downscaling adaptation.
* **Generative Adversarial Networks (GANs)** provide good image quality under moderate noise conditions but degrade with high noise. They handle downscaling better than DIP.
* **SIREN** exhibits impressive visual quality, particularly in retaining fine details. However, its quantitative robustness in noise handling remains to be further tested.