Image enhancement using Deep Learning

Samudrala Hareesh-621243

1. Objective:

Deep learning has revolutionized various fields, including image processing and computer vision. Image enhancement, particularly image super-resolution (SR), aims to improve the quality of images by increasing their resolution. This report details the development and implementation of an Enhanced Super-Resolution Convolutional Neural Network (SRCNN) that utilizes deep learning techniques to upscale low-resolution images while preserving fine details.

2. Theory:

Image Enhancement:

Image enhancement is a key process in digital image processing aimed at improving the visual quality of an image or making certain features more prominent. It doesn't necessarily add new information to the image, but it optimizes the presentation of the existing data for better interpretation by human eyes or computer algorithms.

Convolutional Neural Networks:

Convolutional Neural Networks (CNNs) are a class of deep learning models commonly used for image classification tasks due to their ability to capture spatial hierarchies in images. The key components of CNNs include:

- **Convolutional Layers**: These layers apply filters to input images to extract features like edges, textures, and shapes.
- **Pooling Layers:** These layers reduce the dimensionality of the feature maps, helping the model generalize better and reduce computation.
- **Fully Connected Layers:** At the end of the network, fully connected layers take the flattened feature maps and predict the class probabilities.
- **Activation Functions**: The ReLU activation function introduces non linearity, enabling the network to model complex patterns.
- **Loss Function**: Cross-entropy loss is used to measure the discrepancy between the predicted and true labels, making it ideal for multi-class clas sification tasks.
- **Optimization:** The Adam optimizer is utilized to update the network weights based on gradients, enhancing the learning process.

Super-Resolution Using Convolutional Neural Networks (SRCNN

- **SRCNN** is one of the earliest deep learning-based models for single image super-resolution (SISR).
- **Architecture**: A simple CNN with three layers used to learn mapping from low-resolution to high-resolution images.
- **Advantages**: Fast inference, simple architecture.
- **Disadvantages**: Performance is not as good as more recent models.

3. Methodology

3.1 Dataset Preparation

- **Dataset**: The dataset was sourced from Kaggle, comprising a diverse set of images suitable for super-resolution tasks.
- **Preprocessing**:
    - Images were converted to grayscale for uniformity.
    - High-resolution images were resized to 128x128 pixels for the training dataset.
    - Low-resolution images were generated by further resizing the high-resolution images to 64×64 pixels.

3.2 Enhanced SRCNN Model Architecture

The Enhanced SRCNN model consists of four convolutional layers designed to learn the mapping from low-resolution to high-resolution images:

- **Layer 1**: Convolution with a kernel size 9, followed by ReLU activation and batch normalization.
- **Layer 2**: Convolution with a kernel size 5, followed by ReLU activation and batch normalization.
- **Layer 3**: Convolution with a kernel size 3, followed by ReLU activation and batch normalization.
- **Layer 4**: Convolution with a kernel size 3 and 1 filter (output layer).

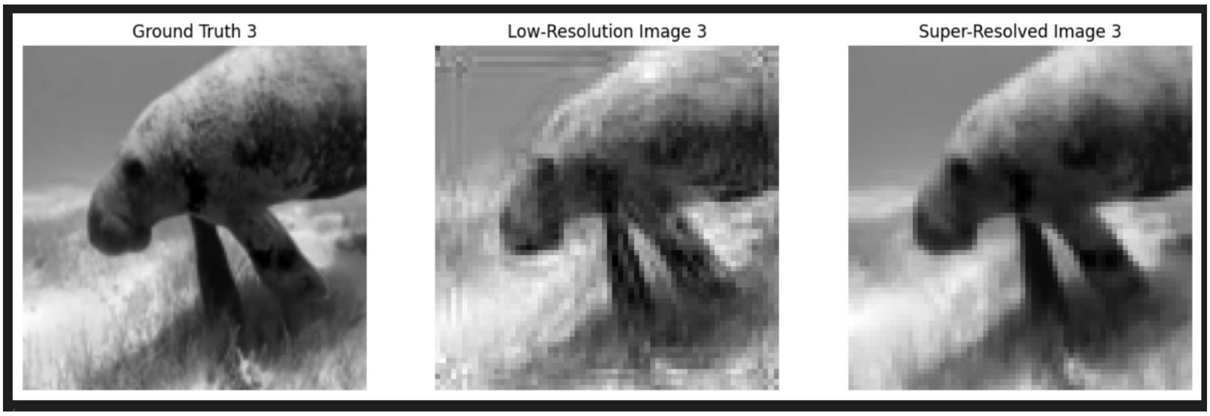
3.3 Training Procedure

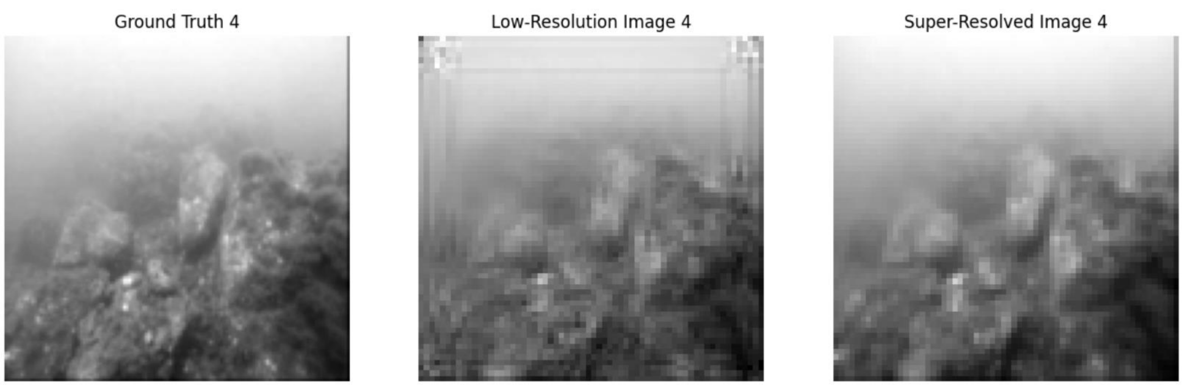
- **Loss Function**: Mean Squared Error (MSE) was employed to quantify the difference between the predicted high-resolution images and the ground truth.
- **Optimizer**: The Adam optimizer was used with a learning rate of 1×10^{-4}
- **Training Configuration**: The model was trained for 10 epochs.

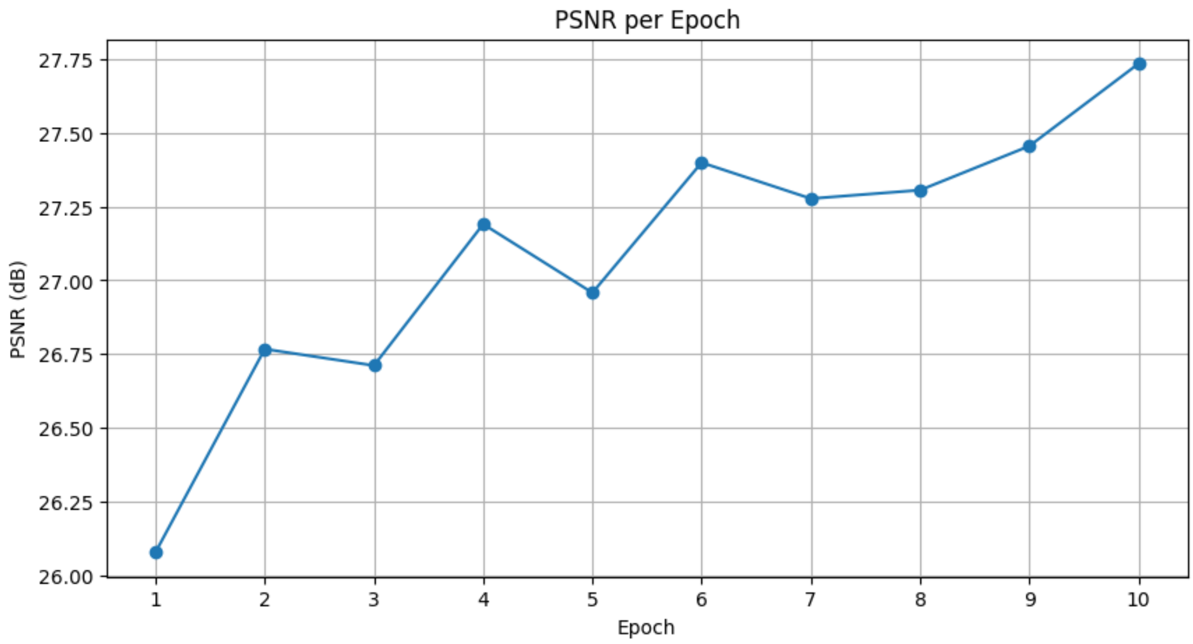
4. Evaluation

To evaluate the performance of these models, we will use:

- **Peak Signal-to-Noise Ratio (PSNR)**: Measures the ratio between the maximum possible value of a signal and the power of corrupting noise. Higher PSNR indicates better quality.
- **Structural Similarity Index (SSIM)**: Measures the similarity between two images. A higher SSIM score indicates better perceptual quality.







The PSNR value starts at around 26 dB in the first epoch and increases steadily, ending at around 27.75 dB by the 10th epoch. This shows that the model is improving its performance as it trains, with the reconstructed images becoming more similar to the reference images over time.

5. Result

- After training the Enhanced SRCNN model, PSNR and SSIM values were calculated on a validation set. The results showed significant improvements in image quality. Hence image enhancement is successfully done using deep learning.

6. References

[1] C. Ancuti, C. O. Ancuti, and P. Bekaert, "Enhancing underwater images and videos by fusion", in Proc. of IEEE Int. Conf. Comput. Vis. Pattern Rec. (CVPR), 2012, pp. 81-88.

[2] C. Li, J. Guo, R. Cong, et al., “Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior,” IEEE Trans. Image Process., vol. 25, no. 12, pp.5664-5677, 2016.

[3] Y. Peng and P. Cosman, “Underwater image restoration based on image blurriness and light absorption”, IEEE Trans. Image Process., vol. 26, no. 4, pp.1579-1594, 2017.

[4] Y. Peng, T. Cao, and P. Cosman, "Generalization of the dark channel prior for single image restoration", IEEE Trans. Image Process., vol. 27, no. 6, pp. 2856-2868, 2018.
C. Li, J. Guo, C. Guo, et al., "A hybrid method for underwater image correction", Pattern Rec. Lett., vol. 94, pp. 62-67, 2017.