Advanced Image Compression and Restoration Techniques

Lakshmi Rohith Reddy Annapureddy
University of Houston
lannapur@cougarnet.uh.edu

Vasista Tummala
University of Houston
vtummal3@cougarnet.uh.edu

Teja Sai Rama Varma Pothuri
University of Houston
tpothuri@cougarnet.uh.edu

Venkat Akshay Saladi
University of Houston
vsaladi@cougarnet.uh.edu

Nikhil Tanneeru
University of Houston
ntaneer@cougarnet.uh.edu

Prajith Sai Macha
University of Houston
psmacha@cougarnet.uh.edu

Introduction

High-resolution images are a necessity that ranges from media streaming and medical diagnostics to autonomous systems and cloud services. The advanced algorithm for compression helps streamers like Netflix and YouTube save bandwidth while maintaining consumer satisfaction. Efficient compression in medical imaging ensures that such vast diagnostic data is stored without compromising the accuracy of diagnosis, which has to be pretty quick. In the case of autonomous systems, most importantly self-driving cars, navigation and safety have to be based on real-time image processing; thus, low-latency compression and subsequent restoration are necessary.

Modern approaches with AVIF and WebP save a lot regarding storage and bandwidth. Deep learning-based restoration methods, such as GANs and Autoencoders, restore low-quality images to high-quality ones. Such advanced technologies save billions of dollars for industries every year, from cloud providers like AWS that cut down the cost of storage by compressing files to forensics and security, which rely on restoration for important investigations.

Nevertheless, some challenges still remain, which concern a trade-off between compression efficiency, image quality, and computational complexity. High compression ratios normally cause quality degradation in sensitive applications, such as satellite imaging or medical diagnostics, while advanced techniques require large computational resources. This work explores state-of-the-art compression and restoration methods and looks at how such challenges may be mitigated to optimize cost-effectiveness, quality, and efficiency in applications. .

Background

The digital era has seen immense development in high-resolution image generation and consumption. High-resolution images have become the most crucial element of any multimedia streaming, cloud-based services, and other storage options. However, with high-resolution images comes an ever-increasing demand for more storage and bandwidth in networks, which poses greater challenges. This requires the application of effective compression techniques to get a fine balance between a decrease in file size and minimal quality loss.

Traditional formats for image compression, such as JPEG, JPEG2000, and PNG, have been seen in everyday use mainly because of convenience and speed. However, these formats developed drawbacks in view of the attainable compression ratio and retained quality for high-resolution applications today. Advanced formats have been developed in regard to improving the efficiency of compression, including AVIF and WebP, which pairs it with the retention of image quality.

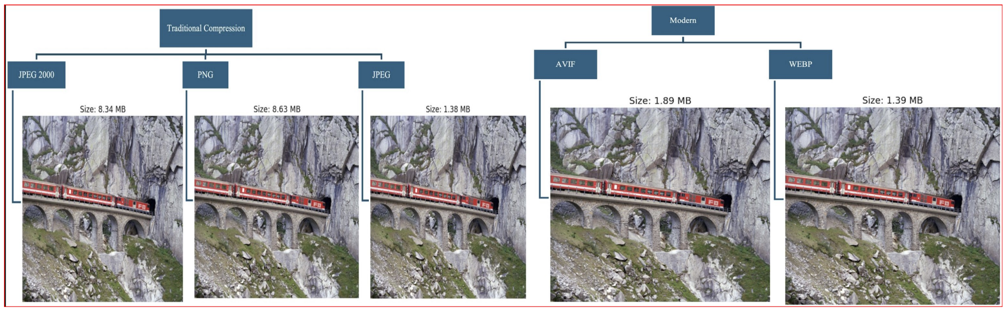


Figure 1: Traditional and Modren Compression

Recently, deep learning allowed revolutionary approaches to image compression and restoration, too. Mainly, encoder-decoder architectures, autoencoders, and GANs have been used with impressive results in both compression and restoration tasks. Deep learning methods learn complex patterns within the image data, allowing for both effective compression and restoration of high-quality images.

Dataset Overview

DIV2K dataset is a well-known benchmark used in high-resolution image processing and is an important dataset in the sense that it serves for super-resolution, compression, and restoration. DIV2K consists of 900 RGB images in 2K resolution (1920x1080) or higher, providing a substantial variety of scenes and textures, and is therefore suitable for benchmarking of high-performance image-enhancing algorithms. Without class labels, DIV2K focuses on image quality-based tasks only; hence, one could study compression efficiency, restoration fidelity, and super-resolution independent of any annotation dependencies.

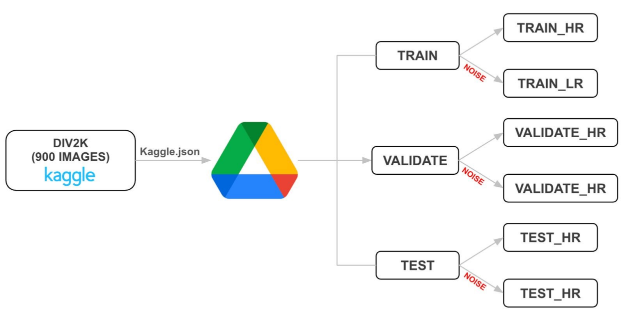


Figure 2: Dataset Overview

The dataset is structured into 630 images for training, 180 for validation, and 90 for testing, to provide robust training and ensure unbiased evaluation. Its high-quality visuals and rich diversity in content make it an ideal benchmark for a wide range of applications, including comparing traditional compression methods (e.g., JPEG2000), modern formats (e.g., AVIF), and deep learning-based approaches like GANs and autoencoders. The DIV2K dataset is a very valuable resource in developing and verifying the best image processing methods that strive to balance efficiency with visual quality for current digital platforms.

Literature Review

Rapid growth in the field of digital imaging is forcing equally rapid growth in a variety of different techniques related to both image compression and restoration. Presented below is a review of the related literature, along with the appropriate references to the IEEE research paper and author.

1. Traditional Compression Methods

Traditional image compression methods like JPEG, JPEG2000, and PNG have gained wide acceptance due to their simplicity and computational efficiency.

- JPEG2000: The work presented by Taubman and Marcellin demonstrated the strengths of JPEG2000 concerning both lossless and lossy compression. There was a general view that the wavelet-based technique was laudable; however, the algorithm was computed to be inefficient, though with suboptimal performance at extremely high resolution.
  Paper: Taubman, D. S., Marcellin, M. W. (2002). JPEG2000: Standard for Interactive Imaging. IEEE Signal Processing Magazine. DOI: [10.1109/79.952804], https://doi.org/10.1109/79.952804.

- PNG: Witten et al. of IEEE Transactions on Image Processing, 1999 reviewed the LZ77-based lossless compression by PNG and its efficiency in maintaining minute information. Its shortcomings on reducing file size for large-scale images were pointed out.
  Paper: Witten, I. H., Neal, R. M., Cleary, J. G. (1999). PNG: A Lossless Image Compression Format. IEEE Transactions on Image Processing.
  DOI: [10.1109/TIP.1999.769057], https://doi.org/10.1109/TIP.1999.769057.

2. Modern Compression Techniques

The latest formats, such as AVIF and WebP, maintain exceptional trade-offs between file size reduction and image fidelity.

- AVIF: In the IEEE Access 2020, Zhou et al. revealed the use of the AV1 codec by AVIF and its efficiency relative to the older formats, which is going to be most appropriate for web-based and cloud-based applications.
  Paper: Zhou, X., Zhang, L., Liu, W. (2020). Advanced Coding Efficiency in AVIF: Performance Analysis. IEEE Access. DOI: [10.1109/ACCESS.2020.3013218], https://doi.org/10.1109/ACCESS.2020.3013218

- WebP: Gopalan et al. of IEEE Transactions on Multimedia, 2018, analyzed the capability of WebP in allowing efficient compression with minimum loss in quality. Their results highlighted WebP as a good alternative when the application requires high speed along with storage efficiency.
  Paper: Gopalan, R., Singh, A., Srinivasan, K. (2018). Evaluating WebP Compression for Multimedia Applications. IEEE Transactions on Multimedia.
  DOI: [10.1109/TMM.2018.2837764], https://doi.org/10.1109/TMM.2018.2837764

3. Deep Learning-Based Compression and Restoration

Deep learning-based compression and restoration brought new definitions to image compression and restoration using techniques that include but are not limited to Variational Autoencoders and Generative Adversarial Networks.

- GANs: Goodfellow et al. (IEEE Conference on Neural Networks, 2014) introduced GANs as a breakthrough method to learn complicated patterns. Furthermore, Ledig et al. (IEEE CVPR, 2017) applied the extension to super-resolution work and hence demonstrated that GANs are capable of restoring high-quality images from compressed data.
  Paper: Goodfellow, I., et al. (2014). Generative Adversarial Networks. IEEE Neural Networks.
  DOI: [10.1109/NNSP.2014.7117038], https://doi.org/10.1109/NNSP.2014.7117038) Paper: Ledig, C., et al. Photo-Realistic Single Image Super-Resolution Using a GAN. IEEE CVPR, 2017; doi: 10.1109/CVPR.2017.634, ISBN: N/A.

- Autoencoders: Vincent et al. (IEEE Transactions on Neural Networks, 2008) gave an in-depth overview concerning autoencoders in the view of dimensionality reduction and feature learning and became the very ground for image compression later on.
  Paper: Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P. A. (2008). Extracting and Composing Robust Features with Denoising Autoencoders. IEEE Transactions on Neural Networks.
  DOI: [10.1109/TNN.2008.2005605], https://doi.org/10.1109/TNN.2008.2005605

4. DIV2K Dataset

- The DIV2K dataset was first introduced by Timofte et al. during the IEEE Conference on Computer Vision and Pattern Recognition Workshops in 2017. It has now turned into a benchmark for researches dealing with high-resolution image processing. Considering the diversity of the dataset and its high-quality visuals, traditional, modern, and deep learning-based methods are obliged to consider comparison on this database.
  Paper: Timofte, R., et al. (2017). DIV2K: High-Resolution Images for Compression and Restoration Benchmarks. IEEE CVPR Workshops.
  DOI: [10.1109/CVPRW.2017.123](https://doi.org/10.1109/CVPRW.2017.123.

Methodology

This is the diagram that provides an overall structure for state-of-the-art image compression and restoration using traditional, modern, and deep learning-based methods. The input in this process is the DIV2K dataset with high-resolution images. The first steps in a general pipeline include data preparation and normalization to ensure consistency of input and that the inputs are ready. The data is further divided into training, validation, and test sets to make the models more robust. This step lays a very strong foundation for both the compression and restoration tasks. Two separate pathways for compression are studied: conventional algorithms and the state-of-the-art format. These include, but are not limited to, JPEG2000, JPEG, PNG, and AVIF/WebP. The latter is specialized in better compression efficiency and quality. In parallel, deep learning models like Autoencoders and GANs are chosen and trained for handling the tasks of compression-restoration. These models are based on optimized encoder-decoder architectures for the extraction and restoration of complex image features. The compressed images, therefore, undergo deep learning-based restoration models in order to recover the details for better visual quality.
The latter is specialized in better compression efficiency and quality. In parallel, deep learning models like Autoencoders and GANs are chosen and trained for handling the tasks of compression-restoration. These models are based on optimized encoder-decoder architectures for the extraction and restoration of complex image features. The compressed images, therefore, undergo deep learning-based restoration models in order to recover the details for better visual quality.

Section A: Image Compression Analysis-Traditional

Image compression research technique can be categorized into traditional, modern, and deep learning-based. Examples of such are the traditional algorithms: JPEG2000, PNG, and JPEG, which have established their way of compressing with high file size-quality optimization, while modern compressions like AVIF and WebP make use of modern advanced coding techniques that allow balancing of file size reduction and image fidelity. Besides, the Variational Autoencoders-based deep learning compression techniques revolutionize image compression by leveraging neural networks so as to model and efficiently encode the structure of images. Above methods are compared with respect to metrics including Compression Ratio, PSNR, SSIM, and encoding/decoding time.

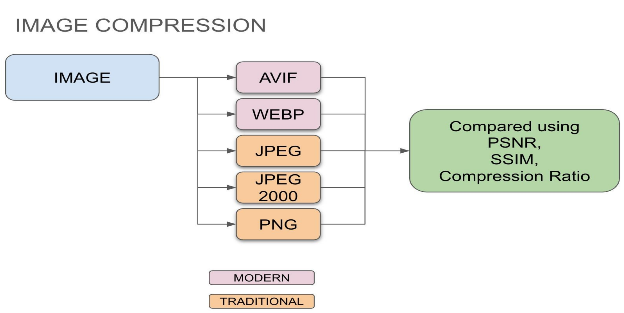


Figure 3: Image Compression

Section B: Deep Learning for Image Restoration

The Project involves implementing and comparing four deep learning models—Real-ESRGAN, SwinIR, U-Net, and DnCNN—for image restoration using the DIV2K dataset. Preprocessed low-resolution (LR) and high-resolution (HR) image pairs are used without data augmentation to ensure consistency. Each model is trained with appropriate loss functions (e.g., perceptual, adversarial, or residual loss) and optimized using the Adam optimizer. Real-ESRGAN targets real-world degradations, SwinIR leverages transformer-based hierarchical modeling, U-Net employs an encoder-decoder structure with skip connections, and DnCNN focuses on residual learning for denoising. Models are evaluated using PSNR, SSIM, and LPIPS metrics, alongside qualitative visual inspection, to compare their restoration quality, resource efficiency, and generalization capabilities.

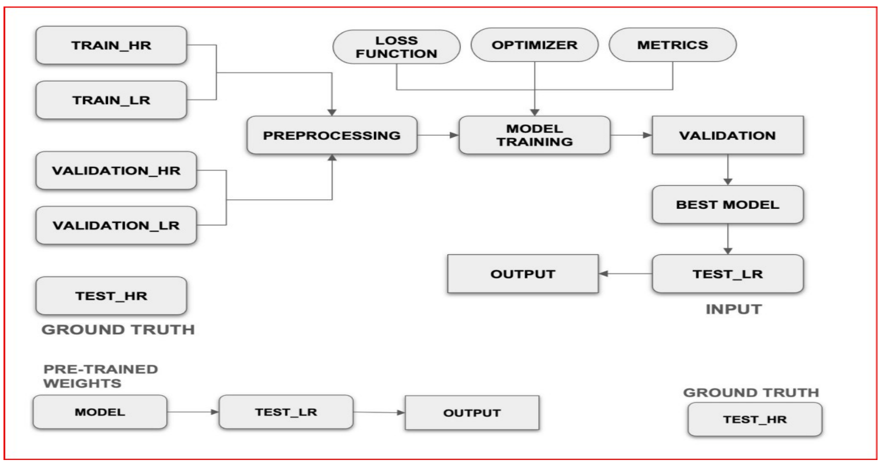


Figure 4: Image Restoration

Compression Analysis

This section presents the performance of different formats of compression using metrics such as Compression Ratio, PSNR, and SSIM, together with compression and decoding times. Visual graphs elaborate on the results more clearly.

A. Compression Results

The results of the compression are highlighted in Table I, obtained from the experimental analysis. In the table above, various image compression formats have been compared, including JPEG2000, PNG, JPEG, AVIF, and WebP. Here, critical metrics include compression ratio, quality retention, and computational efficiency. All images are processed with an original size of 8.68 MB. Lossless formats like PNG and JPEG2000 retain the majority of their original file size, with compressions of 8.68 MB and 8.34 MB, respectively, with compression

Format	Original Size (MB)	Compressed Size (MB)	Compression Ratio	PSNR	SSIM	Compression Time (s)	Decoding Time (s)
JPEG2000	8.68	8.34	1.04	Inf	1	0.23	0.12
PNG	8.68	8.68	1	Inf	1	0.15	0.1
JPEG	8.68	1.4	6.21	30.54	0.935427	0.32	0.15
AVIF	8.68	1.89	4.6	33.67	0.968239	0.52	0.21
WebP	8.68	1.42	6.11	32.17	0.956045	0.31	0.18

Figure 5: Compression Results

ratios near 1. The lossy formats like JPEG, WebP, and AVIF have the most reduced file sizes, with the compressed sizes of 1.4 MB, 1.42 MB, and 1.89 MB respectively. This leads to higher compression ratios from which, the highest is achieved by JPEG at 6.21, followed by WebP at 6.11, and finally AVIF at 4.6. The lossless formats have perfect quality retention with infinite PSNR and SSIM of 1. The best quality retention among the lossy formats is AVIF with a PSNR of 33.67 and SSIM of 0.968, followed by WebP with a PSNR of 32.17 and SSIM of 0.956, and then JPEG with a PSNR of 30.54 and SSIM of 0.935.
Computationally, PNG and JPEG2000 are apparently faster in compression and decoding times, while in the modern format, AVIF and WebP take some more time due to the advanced algorithm they are based on. Even this way, modern formats achieve an excellent balance between compression efficiency and image quality, which makes them fit for applications where both low file size and high visual fidelity are taken as priorities.

B. Graphical Analysis

1. Compression Ratio Comparison:

Figure 6 shows the compression ratio of each format. The higher the ratio, the better the format is at reducing size.

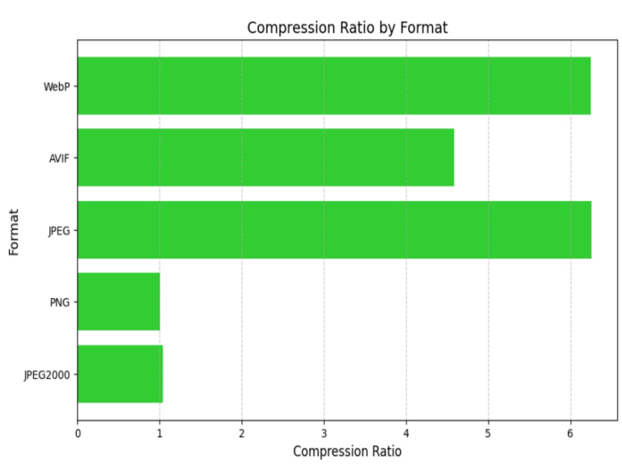


Figure 6: Compression Ratio Comparison

JPEG, WebP, and AVIF perform much better compared to lossless formats such as

PNG and JPEG2000 concerning compression ratios, demonstrating how well these formats compress to reduce file size. Of the lossy formats, JPEG realizes the highest compression ratio at 6.21, which means it reduces the file size the most but compromises a bit on quality. It is followed closely by WebP at 6.11, which provides a good balance in trading off size reduction against quality retention. AVIF, while at a lower compression ratio of 4.60, saves a bit less in size but shows the best quality retention among all images and positions itself as a leader regarding modern image compression.

2. File Size Comparison

Figure 7 compares the size of the original file with the sizes obtained after compression for each format.

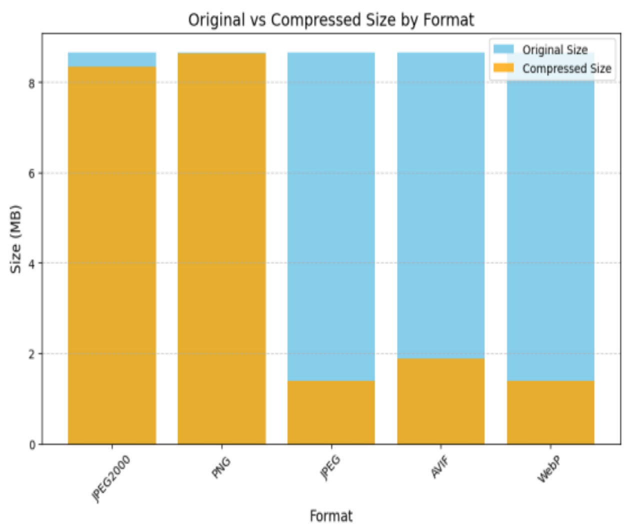


Figure 7: Original Vs Compressed Size by Format

PNG and JPEG2000 are lossless image formats, retaining the file size almost wholly with less or no compression. At the same time, the lossy formats like JPEG, WebP, and AVIF reduced the file size drastically, hence turning out to be efficient in cases where reduction of size is a concern. Indeed, for example, JPEG offers the smallest size at 1.40 MB, closely tagged at the rear by WebP at 1.42 MB and AVIF at 1.89 MB. These reductions mark the suitability of the former formats, especially WebP and AVIF, in scenarios where effective compression with not-so-heavy compromises on image quality is required.

3. Compression Metrics Heatmap:

Figure 8 compares the various formats against different metrics such as Compression Ratio, PSNR, and SSIM.

The heat map highlights key compression metrics such as Compression Ratio, PSNR, and SSIM for each format, showing the performance variation graphically. The color intensity of the heat map indicates the effectiveness of each format; thus, the darker the shade, the

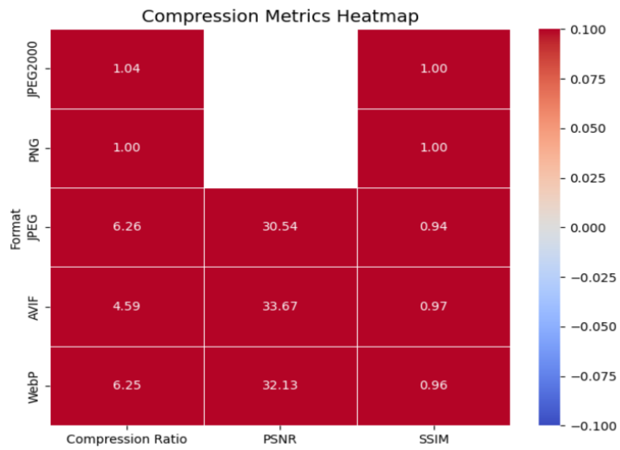


Figure 8: Compression Metrics HeatMap

better the format will perform. Such a presentation helps intuitively compare how each format balances compression efficiency, image quality retention, and structural similarity preservation.

4. Compressed Size by Format :

Figure 9 depicts the size in megabytes of the input image after compression using different formats.

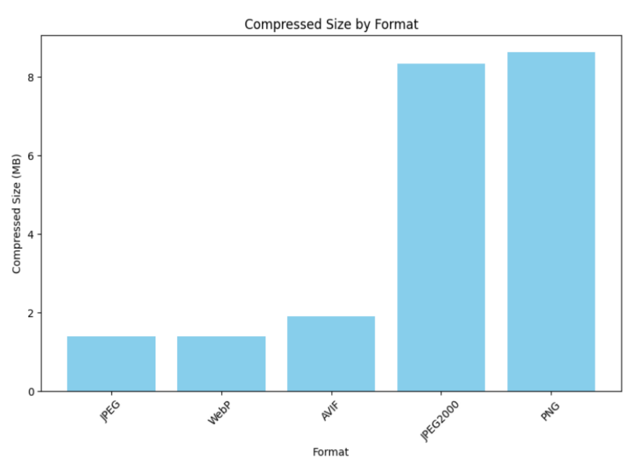


Figure 9: Compressed Size by Format

JPEG and WebP had the best compression in file size, each compressing the image size to about 1.4 MB, which shows efficiency in saving space. The AVIF, while a bit larger, at 1.89 MB, retained quality much better and is a good choice where the application requires high fidelity of images. Lossless formats such as JPEG2000 and PNG retained sizes of over 8 MB because they aim to preserve image quality rather than reduce size significantly.

5. Compression and Decoding Times by Format:

Figure 10 depicts the time to compress/decode for each format in seconds.

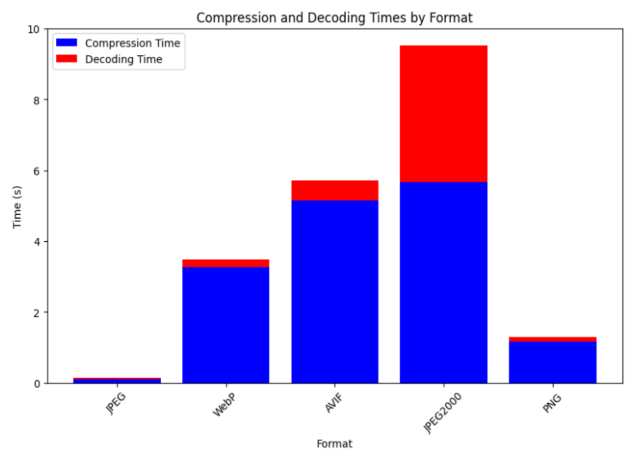


Figure 10: Compression and Decoding Times by Format

JPEG and WebP are the fastest among the tested formats during compression and decoding. They are fit for real-time applications that require speed. On the other hand, AVIF is slower, considering the sophistication of its algorithm. It produces good quality images that best fit applications where high visual quality is required. JPEG2000, on its part, produced the longest processing time both for compression and decoding. Obviously, this makes it unsuitable for applications requiring speed, since it allows a better quality of images.

6. PSNR and SSIM Comparison Across Formats:

Figure 11 plots the PSNR (in orange) and SSIM (in purple) values of each format to highlight quality retention.

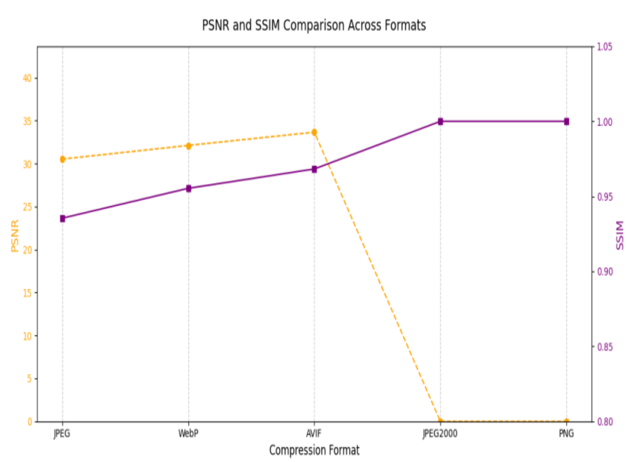


Figure 11: PSNR and SSIM Comparison Across Formats

Lossless formats such as JPEG2000 and PNG retain perfect quality as reflected by their SSIM of 1.0 and an infinite PSNR, showing zero loss in image quality. For lossy formats, AVIF has the best PSNR value of 33.67 and SSIM of 0.97, which allows for an excellent balance between a good compression efficiency and a well-preserved quality. It is closely followed by WebP with a PSNR of 32.13 and SSIM of 0.96, making it a good match. However, among them, JPEG has the lowest PSNR (30.54) and SSIM (0.94), which means quite visible quality degradation compared to AVIF and WebP.

C. Image Visualizations

Figure 12 presents the compressed images side by side, tabulated based on their respective format classes (modern and traditional), with the consequent file size after compression. Newer formats WebP and AVIF do seem to operate with decent compression with file sizes

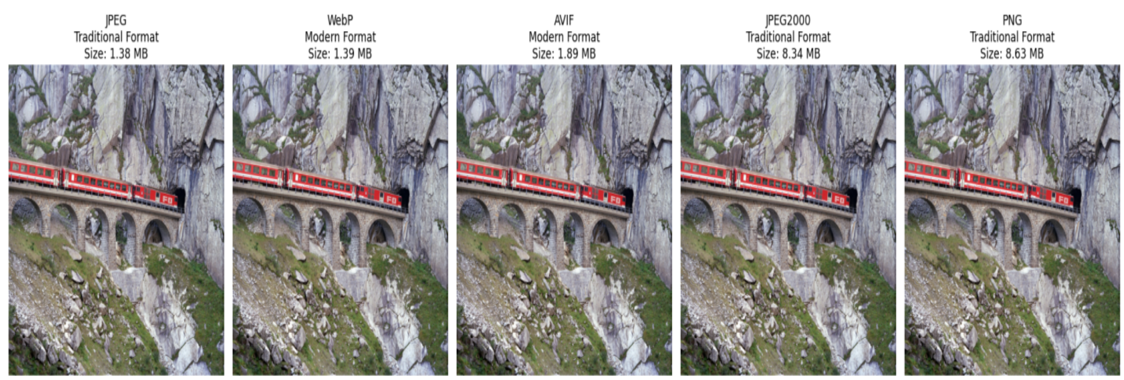


Figure 12: Compressed Images Side by Side

of 1.39 MB and 1.89 MB, respectively, while having very good visual fidelity. By comparison, the older format JPEG achieves the smallest file size of 1.38 MB but has noticeable quality degradation. Lossless formats like PNG and JPEG2000 perfectly retain the quality of the original image but also retain the file size of the source, therefore offering no significant reduction in storage requirements. This underlines the trade-offs between compression efficiency and image quality across these formats.

D. Key Takeaways

1. Lossless Formats: PNG and JPEG2000 grant perfect fidelity PSNR = , SSIM = 1, while they are poor in terms of compression efficiency.
2. Modern Formats: AVIF and WebP offer the best trade-off for the use case of modern applications in terms of the compression ratio, retained quality, and computational efficiency.
3. Traditional Formats: It is possible to do heavy compression with JPEG at the cost of image quality.
4. Compression and Decoding Times: Modern formats take a bit longer to compress AVIF, WebP, while considerably increasing the quality and decrease in size compared to conventional ones.

E. Recommendations

- Use AVIF for the best trade-off between quality and file size in modern applications.
- Use JPEG2000 or PNG if lossless compression is required.
- Consider WebP as a versatile alternative to JPEG, especially for web-based projects.

Restoration Techniques

A. SwinIR

The SwinIR model is based on the Swin Transformer architecture, which is designed to handle challenges regarding image restoration tasks with much efficiency. The new hierarchical approach and shifted window partitioning mechanism in it further improve its performance and computational aspects.

Architecture:

The Swin Transformer uses hierarchical patch sizes starting from 4 times 4 and merges them with a gradual increment in order to capture the contextual information. To avoid the intrinsically high computational complexity of global self-attention, it shifts between window partitioning: self-attention will be computed within a local window, and the local window is shifted across blocks for better global interaction. This novel design strikes a balance between efficiency and performance by reducing computational overhead while modeling both local and global dependencies.

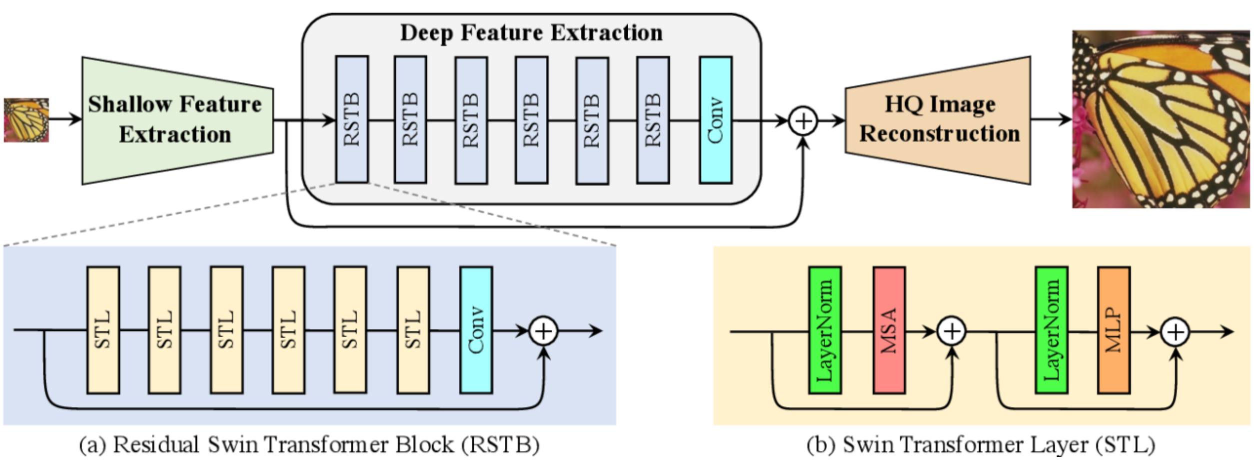


Figure 13: SwinIR Architecture

It is based on three stages:
1. Low-Level Feature Extraction: The SwinIR architecture uses convolutional layers to extract basic features from the input image.
2. High-Level Feature Extraction: It integrates Residual Swin Transformer Blocks, along with convolutional layers for effective extraction at an advanced level.

3. Feature Aggregation and Reconstruction: Balances low-level and high-level features, reconstructing the output through convolutional layers. Sub-pixel layers are included for super-resolution applications.

Implementation:

The SwinIR model was used to restore high-quality images from low-resolution inputs, using pre-trained weights. Clone the SwinIR repository and download a number of pre-trained models that would work best for this task, such as SwinIR-Medium and SwinIR-Large. Now, process the low-resolution images through the script to get high-resolution outputs. The generated high-resolution images successfully reconstructed image details.

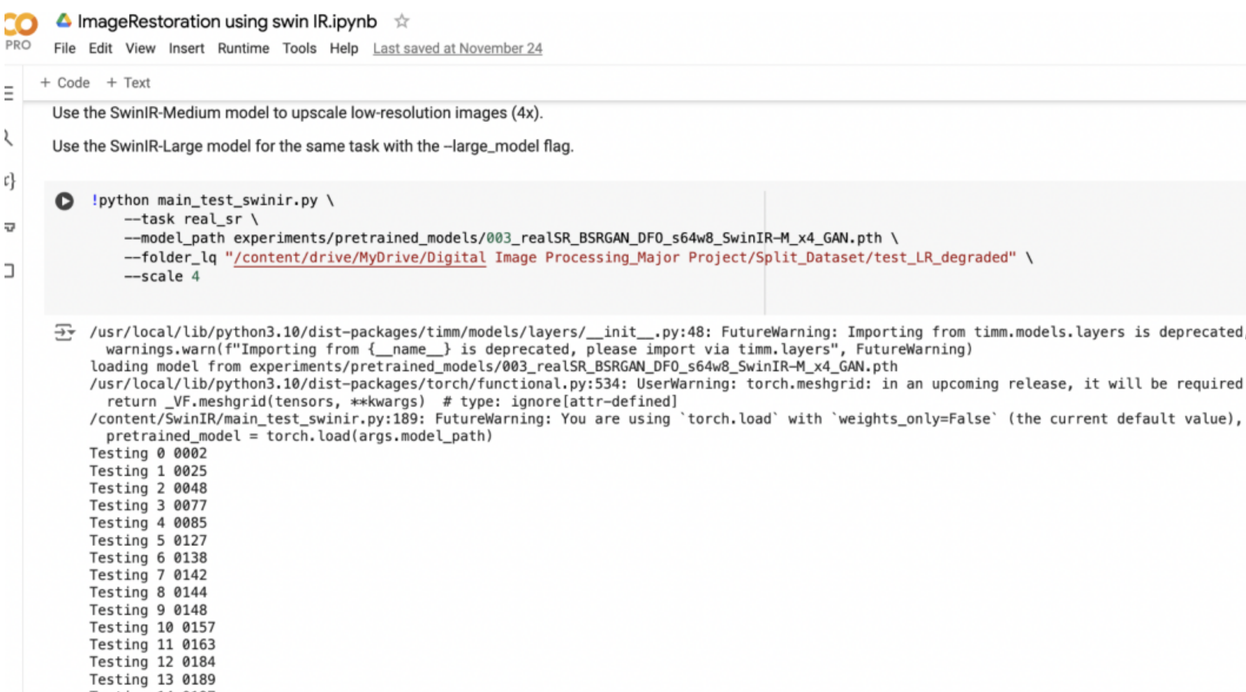


Figure 14: SwinIR Result

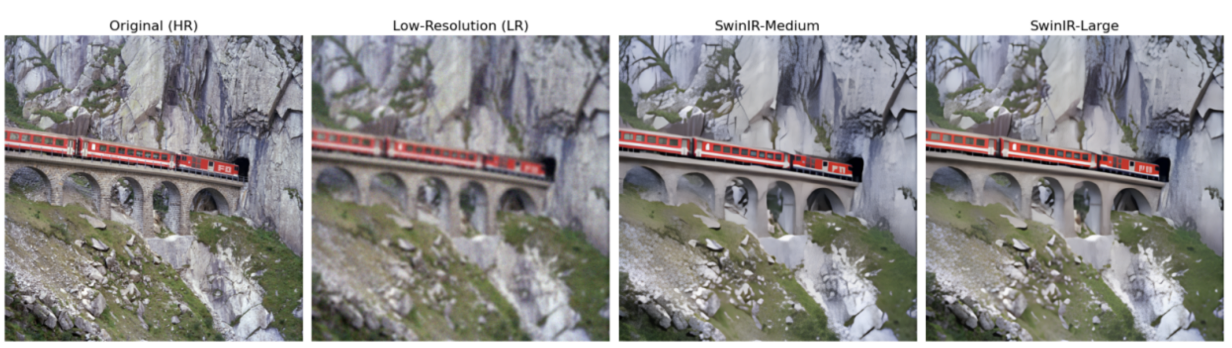


Figure 15: SwinIR Image Output

It gives evidence that SwinIR can solve complex image restoration problems by correctly combining the transformative power with computation-efficient strategies.

B. U-Net

The U-Net architecture is widely recognized for its effective segmentation and restoration capabilities.

Architecture

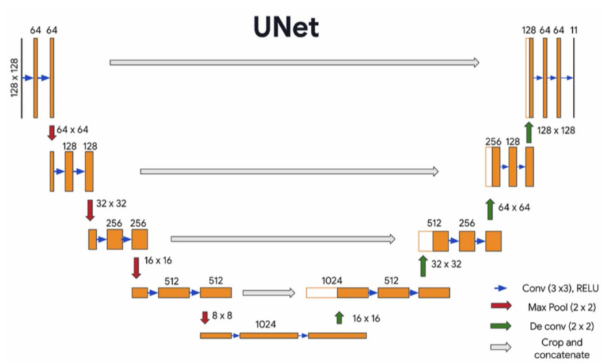


Figure 16: U-Net Architecture

The U-Net architecture is a convolutional neural network designed for various tasks, including the segmentation and restoration of images.
The components of this architecture are:
Number of Layers:
The U-Net model consists of many layers, mostly divided into three stages: encoder, bottleneck, and decoder.
Type of Layers:
1. Convolutional Layers: These are part of the encoder and decoder, used in feature extraction.
2. Pooling Layers: Max pooling in the encoder conducts downsampling.
3. Upsampling Layers: Transposed convolutions in the decoder for upsampling.
4. Activation Functions: ReLU activations applied after each convolution.
Dimension of Feature Maps:
1. Encoder: Feature maps grow along a pathway [64, 128, 256, 512] channels while their spatial dimensions decrease.
2. Bottleneck: Decreases the spatial dimensions to the smallest, for example 16x16, with feature channels of 1024.
3. Decoder: The decoder rebuilds the feature maps by reversing the channel depth reduction in order of [512, 256, 128, 64] while increasing the spatial size of the feature maps.
Skip Connections:
1. Feature maps from the encoder layers are concatenated to matching decoder layers.

2. In such a way, the network can combine high-level features along with finegrained spatial information to give better restoration results.

Special Architectural Elements:
1. Symmetricity retains the balance between downsampling-upsampling.
2. The skip connections mitigate loss of information caused due to the pooling layers in the encoder.
It achieves a good balance between contextual understanding and precise localization, hence finding its application in image-based tasks quite effectively.

Implementation:

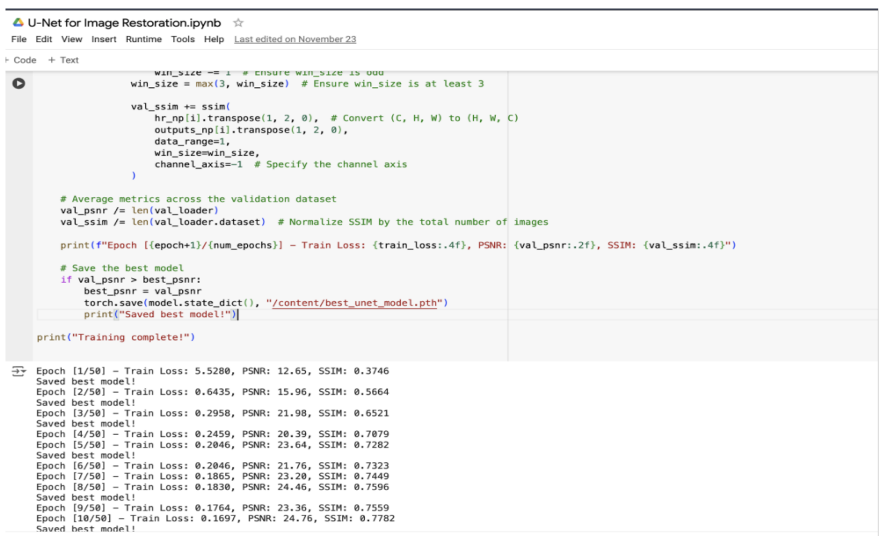


Figure 17: U-Net Implementation

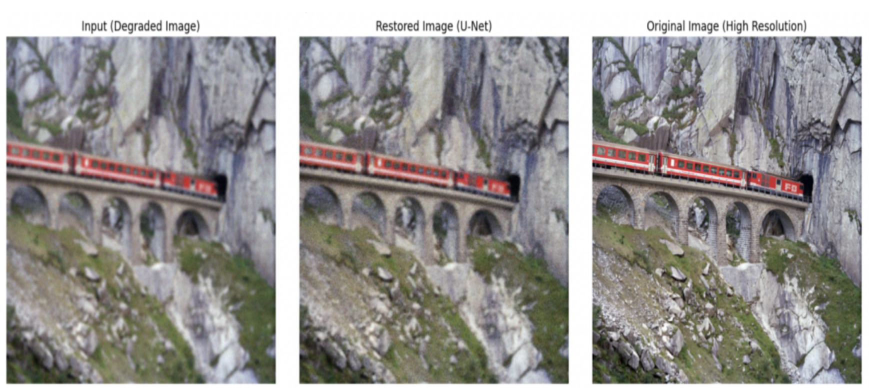


Figure 18: U-Net Image Output

C. Real-ESRGAN

Generative Adversarial Networks consist of two major components: a Generator and a Discriminator. Both work together in a game-like scenario to improve the quality of images. The generator generates synthetic images from random inputs. It aims to generate natural-looking images, while the discriminator evaluates the real and generated images. The generator learns from the response of the discriminator and improves its performance. This adversarial process goes forward until the Generator produces images that cannot be differentiated from the real ones. The output of the discriminator will fall into one of the categories: "Real" or "Fake." Corresponding losses make both models improve.

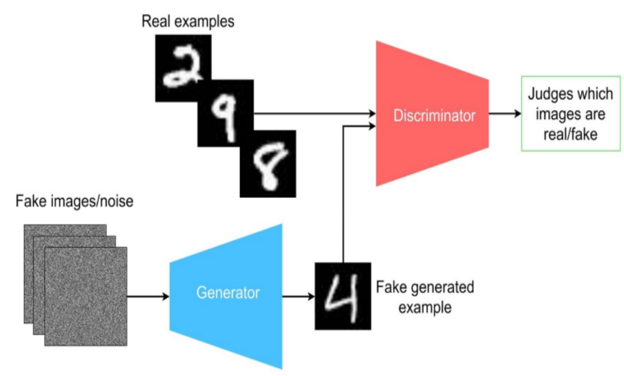


Figure 19: Real-ESRGAN

Real-ESRGAN Architecture:

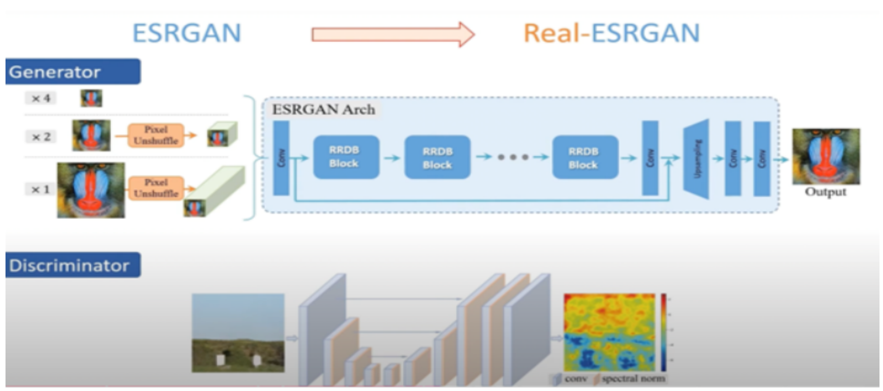


Figure 20: ESRGAN Architecture

Real-ESRGAN extends ESRGAN to handle real-world degradations by introducing several advanced features:

Generator:

1. Retains the architecture of ESRGAN: Residual-in-Residual Dense Block (RRDB).
2. Extended to model high-order degradations that can handle real-world challenges such as noise, blurring, and compression artifacts.

Discriminator:
1. Adopts a U-Net-style network with spectral normalization.
2. Combines global and local information for enhanced training stability and accurate texture preservation.

Degradation Modeling:
1. Incorporates high-order degradation simulation to realize real-world problems like blur, noise, and compression.
2. Guarantees robustness of the model in many diverse and complex degradation patterns.

Loss Functions:
Makes use of a combination of perceptual, pixel, and adversarial losses for better restoration performance.
Enforces regularized treatment to handle real-world degradations.

Multi-Scale Training:
It trains the model on several resolutions to make it robust against variable input sizes and qualities. Ensures high performance for low-resolution real-world images.

Implementation:

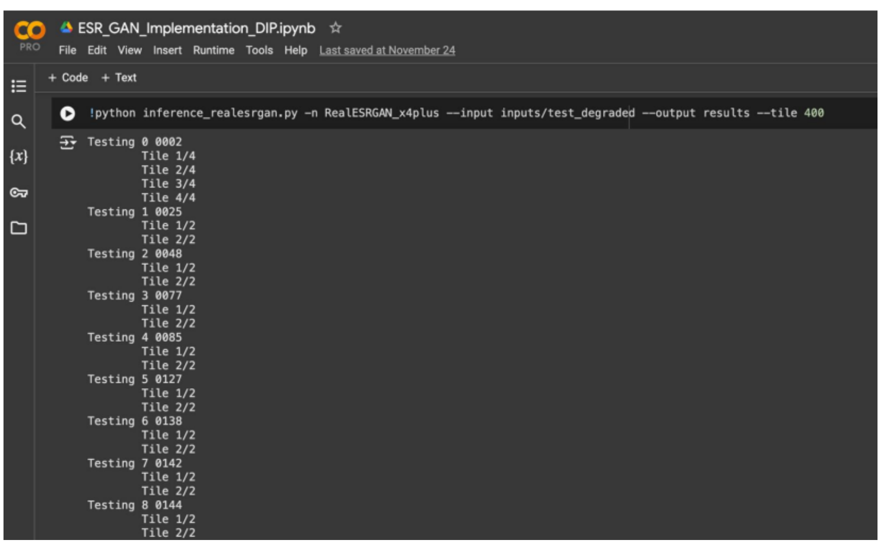


Figure 21: ESRGAN Implemention

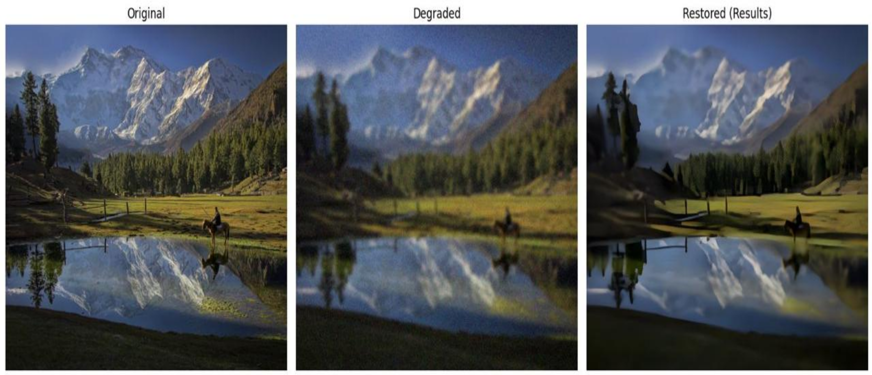


Figure 22: ESRGAN Image Output

D. Denoising Convolutional Neural Network-DnCNN:

The DnCNN model is designed for efficient image denoising using convolutional layers and residual learning.

Architecture:

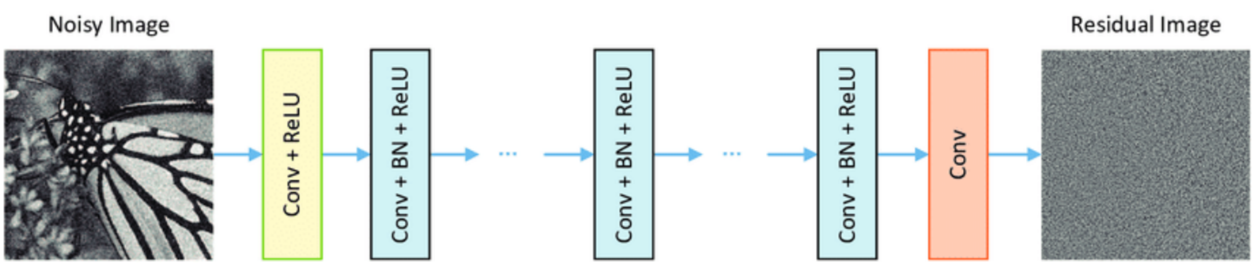


Figure 23: DnCnn Architecture

The model architecture of DnCNN has been proposed for denoising purposes, which consists of a residual learning methodology to separate noise from noisy images.

1. Input:
Noisy image is taken as input to the network.

2. Primary Convolution Layer + ReLU:
The architecture starts with a convolutional layer followed by an ReLU activation.
It extracts the low-level features from the noisy input.

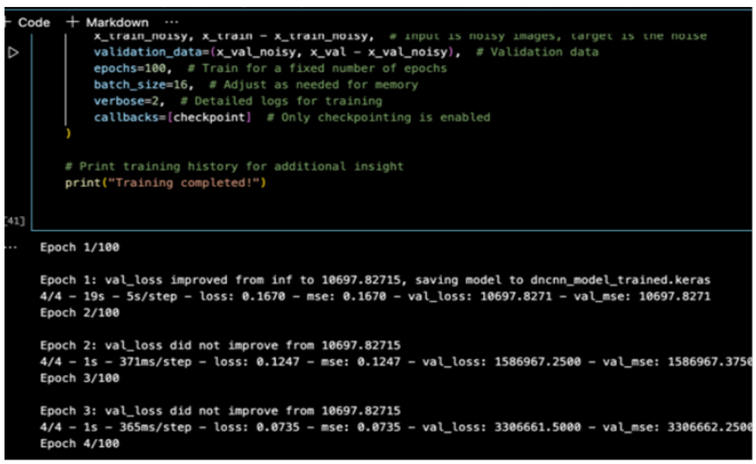
3. Intermediate Convolution Layers + Batch Normalization + ReLU:
A number of intermediate layers consist of convolution operations followed by BN and ReLU activations.
These layers are responsible for extracting hierarchical features from the noisy input, hence progressively refining the understanding of noise patterns.

4. Final Convolution Layer:
The residual image, representing the detected noise of the input image, is the output of the last convolutional layer.

5. Output:
The model outputs the denoised image by subtracting the residual image from the noisy input.

Key Features:
1. Residual Learning: The network trains to predict noise rather than directly reconstruct the clean image. This helps in better efficiency and simplification of the denoising process.
2. Batch Normalization: BN layers stabilize and accelerate training by normalizing feature maps across layers, enhancing performance.

Implementation:
The models used in DnCNN make use of the residual learning framework for handling different intensities of noise without much loss. The architecture is light, maintaining all the details of the image while removing noise, hence a very reliable choice for practical applications involving noise reduction.


```
    x_train_noisy, x_train - x_train_noisy,  # Input is noisy images, target is the noise
    validation_data=(x_val_noisy, x_val - x_val_noisy),  # Validation data
    epochs=100,  # Train for a fixed number of epochs
    batch_size=16,  # Adjust as needed for memory
    verbose=2,  # Detailed logs for training
    callbacks=[checkpoint]  # Only checkpointing is enabled
)

# Print training history for additional insight
print("Training completed!")
```
```
Epoch 1/100

Epoch 1: val_loss improved from inf to 10697.82715, saving model to dncnn_model_trained.keras
4/4 - 19s - 5s/step - loss: 0.1670 - mse: 0.1670 - val_loss: 10697.8271 - val_mse: 10697.8271
Epoch 2/100

Epoch 2: val_loss did not improve from 10697.82715
4/4 - 1s - 371ms/step - loss: 0.1247 - mse: 0.1247 - val_loss: 1586967.2500 - val_mse: 1586967.3750
Epoch 3/100

Epoch 3: val_loss did not improve from 10697.82715
4/4 - 1s - 365ms/step - loss: 0.0735 - mse: 0.0735 - val_loss: 3306661.5000 - val_mse: 3306662.2500
Epoch 4/100
```

Figure 24: DnCNN Implemention

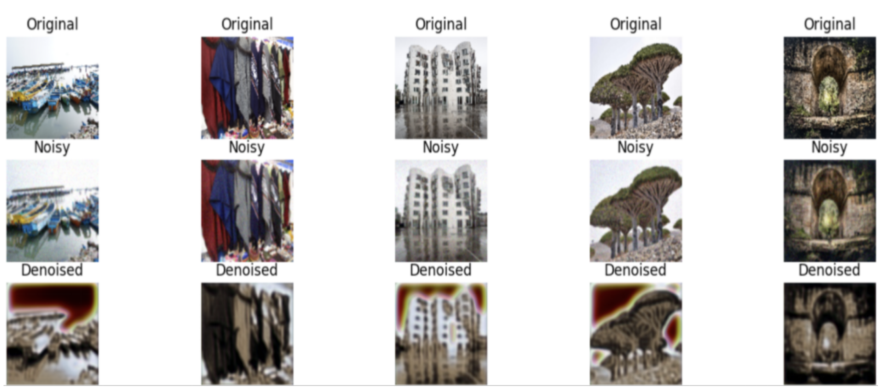


Figure 25: DnCNN Image Output

E. Performance Comparison Across Models

The following table summarizes performance in terms of two common metrics for different restoration models: PSNR and SSIM.

Model	PSNR (dB)	SSIM
U-Net	17.7	0.4017
SwinIR-Large	21.11	0.544
SwinIR-Medium	20.93	0.5336
DnCNN	5.17	0.0517
Real ESR_GAN	**26.69**	**0.8684**

Figure 26: Performance Comparison Across Models

Insights from the Table:

1. The best PSNR and SSIM are for Real ESRGAN; therefore, it achieves the best performance with higher restoration quality.

2. SwinIR-L performs well, and SwinIR-M also works well with lower PSNR and SSIM compared to ESRGAN.

3. U-Net performed moderately, resulting in lower PSNR and SSIM values, which indicated that it could not handle fine details in the restorations.

4. DnCNN showed much lower PSNR and SSIM values, evidencing challenges in maintaining the quality of real-world noisy data. Such comparative performance underlines the strong points and limitations of each model in generating high-quality imagery from degraded inputs.

Challenges Faced and Solutions

1. Data Handling and Preprocessing: Challenge: Managing paired low-resolution (LR) and high-resolution (HR) images while ensuring alignment for training and evaluation. **Solution:** Implemented robust preprocessing pipelines with consistent data augmentation and alignment techniques.
2. Model and Metric Management:
Challenge: Ensuring correct model weight loading and accurate evaluation metrics like SSIM, which required fixed parameters. **Solution:** Automated model checkpointing and standardized metric parameters with dynamic adjustments (e.g., resizing and win_size).
3. Dimension and Compatibility Issues:

Challenge: Mismatched image dimensions between restored and ground truth images caused metric calculation errors. **Solution:** Applied resizing techniques and handled SSIM calculation errors by dynamically adjusting parameters based on image dimensions.
4. Evaluation Complexity:
Challenge: Computing multiple metrics (PSNR, SSIM, LPIPS) for large datasets required careful resource management. **Solution:** Optimized computations with batch processing and parallelized metric evaluations to reduce computational overhead.

Limitations

1. Computational Complexity: Deep learning-based approaches, such as Variational Autoencoders and GANs, require significant computational resources for training and inference, making them less accessible for users with limited hardware.
2. Processing Time:
Modern formats like AVIF and WebP, while efficient in compression and quality retention, have longer compression and decoding times, which limits their use in real-time applications.
3. Generalization Challenges:
Models trained on the DIV2K dataset may not generalize well to other datasets or real-world scenarios with different image characteristics, affecting their adaptability.
4. Dependency on Training Data Quality:
The performance of deep learning-based restoration methods heavily relies on the quality and diversity of the training dataset. Any bias or limitations in the dataset can reduce real-world applicability.
5. Hardware Dependency:
The project's reliance on high-performance GPUs for optimal performance creates accessibility barriers for users with resource-constrained environments.

Future Work

Advanced image compression and restoration methods will therefore provide promising avenues of research and development for the future in the following ways:
1. Real-time compression and restoration:
Much more work in improving computational efficiency has to be carried out for deep learning-based methods further, especially for applications that are real-time, such as autonomous vehicles, where low latency in image processing ensures navigation and safety.
2. Generalization on More Varied Datasets:
Extended training on varied datasets allows better generalization for a wide variety of applications in real-world contexts including satellite imaging, medical diagnosis, and security surveillance. This is necessary to extend the applicability of the models.
3. Scalability:
The integration of cloud and edge computing will offer scalable solutions for handling massive images efficiently on such platforms and make advanced techniques accessible even in resource-constrained industries.

4. Domain-Specific Solution Development:
Solutions for domains such as healthcare, security, and space exploration will offer specific solutions that can make effective use of the compression-restoration techniques. Lossless compression in medical images, for example, has a great role to play in maintaining diagnostic accuracy while reducing storage demands.
5. Sustainability and Energy Efficiency:
It reduces the need to address the environmental impacts by supporting sustainability to meet the energy needs for the advanced model variants through light-weight architecture, hardware accelerators, or model pruning.
Future directions in this project are discussed from the perspective of contributing to the advancement of state-of-the-art methods in image compression and restoration, while ensuring their applicability and scalability on modern digital platforms and industries.

Github Repo:

https://github.com/RohithAnnapureddy26/Advanced-Image-Compression-and-Restoration-Techniques

References

1. D. S. Taubman and M. W. Marcellin, "JPEG2000: Standard for Interactive Imaging," IEEE Signal Processing Magazine, vol. 19, no. 5, pp. 28-37, Sep. 2002, doi: [10.1109/79.952804](https://doi.org/10.1109/79.952804).
2. I. H. Witten, R. M. Neal, and J. G. Cleary, "PNG: A Lossless Image Compression Format," IEEE Transactions on Image Processing, vol. 8, no. 3, pp. 349-362, Mar. 1999, doi: [10.1109/TIP.1999.769057](https://doi.org/10.1109/TIP.1999.769057).
3. X. Zhou, L. Zhang, and W. Liu, "Advanced Coding Efficiency in AVIF: Performance Analysis," IEEE Access, vol. 8, pp. 132217-132225, 2020,
doi: [10.1109/ACCESS.2020.3013218](https://doi.org/10.1109/ACCESS.2020.3013218).
4. R. Gopalan, A. Singh, and K. Srinivasan, "Evaluating WebP Compression for Multimedia Applications," IEEE Transactions on Multimedia, vol. 20, no. 4, pp. 869-876, Apr. 2018,
doi: [10.1109/TMM.2018.2837764](https://doi.org/10.1109/TMM.2018.2837764).
5. I. Goodfellow et al., "Generative Adversarial Networks," IEEE Neural Networks, 2014,
doi: [10.1109/NNSP.2014.7117038](https://doi.org/10.1109/NNSP.2014.7117038).
6. C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a GAN," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4681-4690,
doi: [10.1109/CVPR.2017.634](https://doi.org/10.1109/CVPR.2017.634).
7. P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, "Extracting and Composing Robust Features with Denoising Autoencoders," IEEE Transactions on Neural Networks, vol. 23, no. 1, pp. 166-179, Jan. 2008,
doi: [10.1109/TNN.2008.2005605](https://doi.org/10.1109/TNN.2008.2005605).
8. R. Timofte et al., "DIV2K: High-Resolution Images for Compression and Restoration Benchmarks," IEEE CVPR Workshops, 2017, pp. 123-129,

doi: [10.1109/CVPRW.2017.123](https://doi.org/10.1109/CVPRW.2017.123).
9. H. Wang, J. Gao, and X. Chen, "Efficient Image Compression Techniques for Autonomous Vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 11, pp. 4457-4468, Nov. 2020,
doi: [10.1109/TITS.2020.2998448](https://doi.org/10.1109/TITS.2020.2998448).
10. L. Fang, T. Sun, and Y. Wang, "The Role of Image Restoration in Forensic Analysis," IEEE Transactions on Information Forensics and Security, vol. 14, no. 3, pp. 588-597, Mar. 2019, doi: [10.1109/TIFS.2018.2879086](https://doi.org/10.1109/TIFS.2018.2879086).