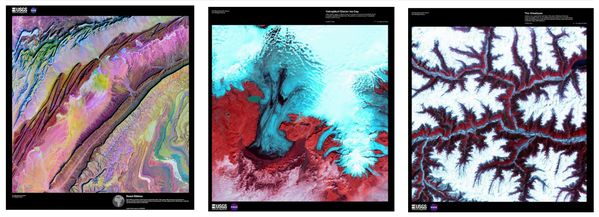
The Art of Compression: JPEG and JPEG-2000 on Multispectral Satellite Imaging

Kritika Pandit, Lucas Schattenmann, Anika Rajbhandary, Leon Liang, Advised by Layla Oesper

INTRODUCTION

➡Satellite images are large and expensive to store/ transmit due to their **high spatial** and **spectral detail**.
➡Lossy compression reduces file size by removing visually insignificant information while maintaining perceptual quality.
➡**JPEG** and **JPEG-2000** use different transforms and coding methods to achieve efficient compression.

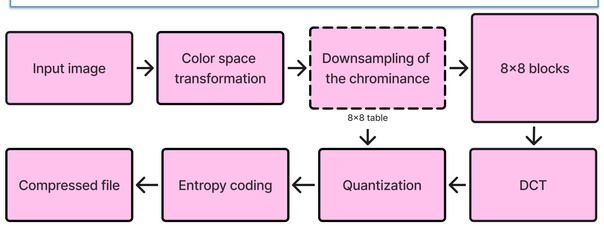


OBJECTIVES

➡ Build a complete JPEG/JPEG 2000 encoder-decoder from scratch to visualize how each compression stage affects image quality.

MATERIALS AND DATASETS

➡**Dataset:** Satellite TIFF images (Earth as Art), chosen for their large file sizes and rich spatial/color detail → Realistic test for lossy compression.
➡**JPEG (from scratch):** Python implementation of full pipeline (RGB→YCbCr, chroma Downsampling, 2D DCT, quantization, zig-zag, RLE, Huffman, bitstream) + Reverse for the Decoder.
➡**JPEG (library baseline):** Pillow-based JPEG encoder used to see behavior and compare metrics against the custom implementation.
➡**Analysis stack:** NumPy, OpenCV, SciPy, scikit-image, and to compute metrics and generate plots.



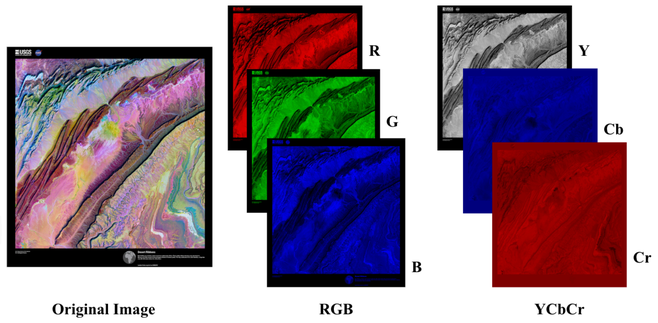
METHODS

JPEG pipeline with a library

➡Python pipeline using **Pillow** (JPEG) and **imagecodecs** (JPEG-2000) + Calculated the Metrics.

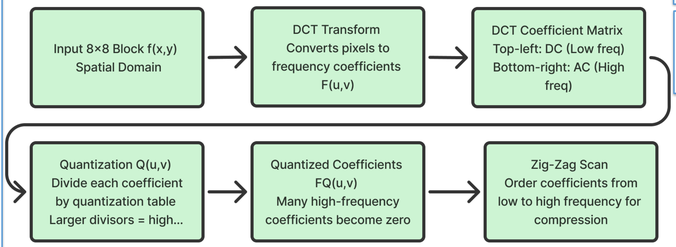
JPEG Pipeline from Scratch

Convert input images from RGB → YCbCr to separate brightness (Y) from color information (Cb, Cr).

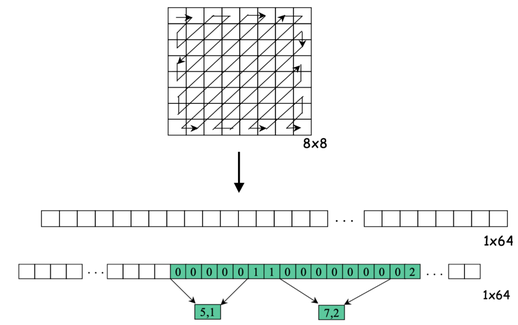


Chroma subsampling reduces the **color resolution** but preserves **luminance detail**, using the fact that the human eye is far less sensitive to color changes than to brightness variations.

➡JPEG compresses an image block-by-block.
➡Each 8×8 region is transformed from pixels into frequency components using the **Discrete Cosine Transform (DCT)**.
➡It is then **quantized** to remove subtle high-frequency details that the human eye barely notices.

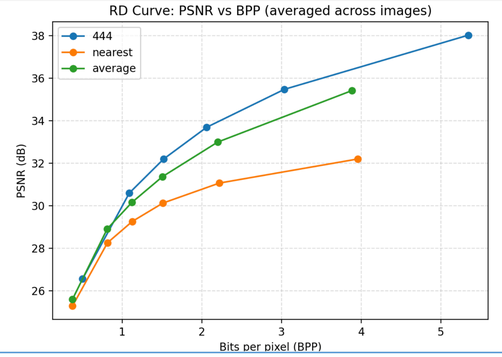


➡After quantization, the 8×8 block is reordered using a **zig-zag scan**, which arranges low-frequency coefficients first and high-frequency ones later.
➡This ordering groups long sequences of zeros together, making the data easier to compress.
➡Next, **Run-Length Encoding (RLE)** replaces these long zero sequences with shorter pairs of numbers, reducing redundancy.
➡Finally, **Huffman coding** assigns shorter binary codes to frequent values (like zeros) and longer codes to rare values, producing the final entropy-coded JPEG bitstream.



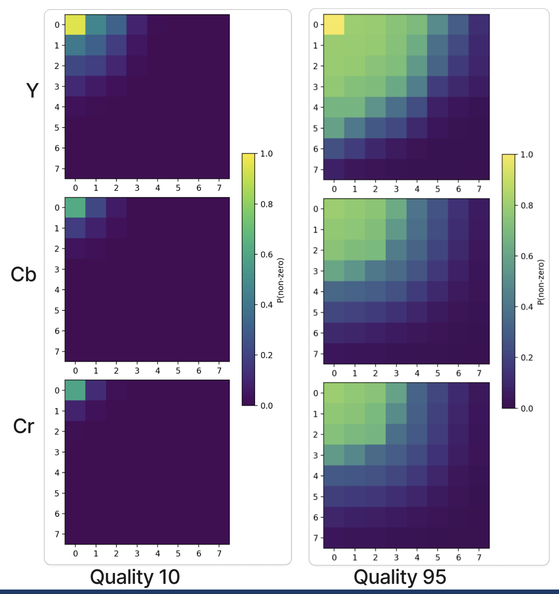
RESULTS

Rate- Distortion Curves to show the trade-off between **compression** (BPP) and **image quality** (PSNR/SSIM).

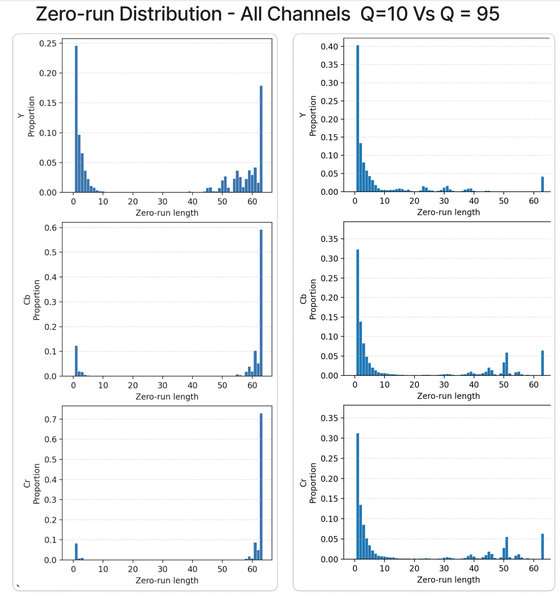


➡This plot shows how **image quality** (measured by PSNR, or Peak Signal-to-Noise Ratio) improves as **more bits** are used per pixel (less compressed).

Coefficient Occupancy Heatmaps to find where **frequency energy** is concentrated and how **sparsity varies**.



Zero-Run Length Distributions to see how quantization and subsampling create long zero runs that improve RLE efficiency across channels.



➡At lower quality (Q=10), quantization creates l**onger zero runs**, especially in chroma channels, making the image more compressible, while higher quality (Q=95) **retains more non-zero coefficients**, leading to less compression.

FUTURE STEPS

➡**ROI-based compression** can be used to preserve important regions, such as land or urban areas, while applying stronger compression elsewhere to save space.
➡**Machine learning** can further enhance this by automatically detecting key regions and adjusting compression levels adaptively.

ACKNOWLEDGEMENTS

I would like to thank **Layla** for her continuous guidance and feedback throughout this project. I am grateful to **Mike Tie** for arranging the computers. I want to thank my **teammates** for their collaboration, support & dedication.

WORKS CITED

1. *Earth As Art | U.S. Geological Survey*. https://www.usgs.gov/centers/eros/earth-art. Accessed 9 Nov. 2025.
2. Luo, Xiyang, et al. "The Rate-Distortion-Accuracy Tradeoff: JPEG Case Study." arXiv:2008.00605, arXiv, 3 Aug. 2020. *arXiv.org*, https://doi.org/10.48550/arXiv.2008.00605.
3. Hosseini, Mohammad. "A Survey of Data Compression Algorithms and Their Applications." arXiv:2506.10000, arXiv, 20 Feb. 2025. *arXiv.org*, https://doi.org/10.48550/arXiv.2506.10000.
4. Khan, Sulaiman, et al. "An Efficient JPEG Image Compression Based on Haar Wavelet Transform, Discrete Cosine Transform, and Run Length Encoding Techniques for Advanced Manufacturing Processes." *Measurement and Control*, vol. 52, nos. 9–10, Nov. 2019, pp. 1532–44. *SAGE Journals*, https://doi.org/10.1177/0020294019877508.