DSP-Aware Convolutional Autoencoder for Satellite Image Compression in 5G Network Slicing

Project Report

Submitted by:

Om Singh (23105131)
Waris (23105074)
Arkaprabha (23105077)

Supervisors:

Dr. Deepak Bagai
Dr. Ashish Singh

Electronics and Communications Engineering Department
Punjab Engineering College (Deemed to be University), Chandigarh

Academic Session: 25261

DECLARATION

We hereby declare that the project work titled **“DSP-Aware Convolutional Autoencoder for Satellite Image Compression in 5G Network Slicing”** is an original and authentic record of our work carried out at Punjab Engineering College (Deemed to be University), Chandigarh.

This project is being submitted as part of the requirements for the award of the degree of **Bachelor of Technology** in **Electronics and Communication Engineering** for the course **Minor Project - 1 (ECP 5101)**. The work has been conducted under the guidance of **Dr. Deepak Bagai** and **Dr. Ashish Singh**.

Date: December 12, 2025
Place: Chandigarh

Om Singh	**Waris**	**Arkaprabha**
(23105131)	(23105074)	(23105077)

ACKNOWLEDGEMENT

We would like to extend our heartfelt gratitude to everyone who contributed to the success of our project, entitled **“DSP-Aware Convolutional Autoencoder for Satellite Image Compression in 5G Network Slicing.”**

First and foremost, we express our sincere thanks to our project guides, **Dr. Deepak Bagai** and **Dr. Ashish Singh**, for their invaluable mentorship, patience, and technical support throughout this research. Their guidance was instrumental in navigating the complexities of integrating Deep Learning with 5G Telecommunications.

We also extend our gratitude to the Coordinator and the Head of the Electronics and Communication Engineering Department, as well as the Punjab Engineering College Administration, for providing the necessary computational resources and facilities that facilitated this project.

Special thanks go to our families for their constant support and encouragement throughout this journey. Their belief in us served as a vital source of motivation and strength. Lastly, we would like to acknowledge the individuals, organizations, and open-source communities whose resources we consulted during this project. Their contributions significantly enriched our understanding of the subject matter and helped us shape a comprehensive and insightful analysis.

Om Singh (23105131)
Waris (23105074)
Arkaprabha (23105077)

Contents

Abstract

Standard image compression standards, particularly JPEG, rely heavily on the Discrete Cosine Transform (DCT). While efficient for general photography, this mathematical approach aggressively suppresses high-frequency components, leading to a significant loss of scientific fidelity in satellite imagery where pixel-level detail is critical. This project presents the design and implementation of a **DSP-Aware Convolutional Autoencoder** that overcomes these limitations by learning non-linear transforms tailored specifically for remote sensing data.

The system utilizes a **Hyperprior architecture** that separates image data into a bulk latent vector (y) and a probabilistic entropy key (z). To ensure high fidelity, the model was trained using a custom frequency-aware loss function that integrates Perceptual Loss (VGG) to preserve structural integrity often discarded by standard codecs. The development journey evolved from a basic autoencoder, which yielded suboptimal results, to a sophisticated model incorporating Generalized Divisive Normalization (GDN) and attention mechanisms.

Furthermore, the project bridges the gap between AI and Telecommunications by integrating **5G Network Slicing**. Using MATLAB's 5G and Satellite Toolboxes, we simulated a transmission pipeline where the bulk image data is routed through high-throughput eMBB slices, while the critical decoding keys are secured via ultra-reliable URLLC slices. The entire system is orchestrated via a Streamlit dashboard running in a WSL (Windows Subsystem for Linux) environment, demonstrating a complete end-to-end workflow from image capture to ground station reconstruction. Quantitative evaluations confirm that the proposed system achieves superior PSNR and SSIM metrics compared to JPEG, validating its potential for next-generation satellite telemetry.

Chapter 1

Introduction

1.1 Background

The volume of data generated by modern earth observation satellites has exploded in recent years, driven by sensors capable of higher spatial and spectral resolutions. Transmitting this massive amount of data from orbit to ground stations presents a significant bottleneck due to limited bandwidth and short communication windows. Traditionally, this challenge has been met using mathematical compression standards like JPEG and JPEG 2000. While these standards are mathematically efficient, they are "content-agnostic"—they treat a critical satellite feature (like a vehicle or crop line) with the same compression logic as a patch of clouds, often leading to the loss of scientifically vital information.

1.2 Problem Statement

The fundamental limitation of current satellite transmission lies in the **Discrete Cosine Transform (DCT)**, the backbone of JPEG compression. Our initial analysis revealed that DCT functions by converting spatial pixel data into frequency components and aggressively quantizing the high-frequency coefficients to save space.

In standard photography, this is acceptable. However, in satellite imagery, high-frequency components represent the sharp edges, textures, and small objects that define the image's value. By suppressing these frequencies, standard codecs introduce "blocking artifacts" and blurring, rendering the images less useful for downstream scientific analysis. There is a critical need for a compression system that is "DSP-Aware"—one that understands the frequency domain implications of compression and preserves the structural integrity of the image.

1.3 Objectives

The primary goal of this project is to develop a deep learning-based solution that outperforms traditional mathematical codecs in high-fidelity satellite image compression. The specific objectives are:

- To design and develop a **Convolutional Autoencoder** capable of learning non-linear transforms for satellite data.
- To implement a sophisticated **Hyperprior Architecture** that separates image data into latent vectors and entropy models for efficient encoding.
- To integrate a **DSP-Aware Loss Function** utilizing Perceptual Loss (VGG) to guide the AI in preserving high-frequency details.
- To simulate a **5G Network Slicing** transmission environment using MATLAB, ensuring reliable delivery of critical decoding keys (URLLC) and high-throughput transmission of image data (eMBB).
- To create a unified **Ground Station Dashboard** (Streamlit) that demonstrates the end-to-end pipeline from encoding to reconstruction.

1.4 Scope

This project focuses on the compression of RGB satellite imagery, utilizing the EuroSAT dataset for benchmarking. The AI model development is conducted in Python (PyTorch), while the telecommunications simulation leverages MATLAB's 5G and Satellite Toolboxes. The system integration is demonstrated within a Windows Subsystem for Linux (WSL) environment to overcome cross-platform compatibility issues encountered during development. The scope is limited to still image compression and does not extend to real-time video stream processing or onboard hardware implementation (FPGA).

Chapter 2

Literature Review

2.1 Standard Image Compression and its Limitations

The dominant standard for digital image compression remains the Joint Photographic Experts Group (JPEG) format. The core engine of JPEG is the **Discrete Cosine Transform (DCT)**, a mathematical operation that converts spatial pixel data into the frequency domain.

The fundamental principle driving DCT-based compression is ”energy compaction.” In typical photography, the majority of an image's energy is concentrated in low-frequency coefficients. JPEG exploits this by using a quantization table to aggressively suppress high-frequency coefficients, which are assumed to be perceptually insignificant.

However, satellite imagery differs fundamentally from natural photography. It is characterized by high entropy and critical ”pixel-sized details”—such as vehicles, road markings, or crop lines—that manifest as high-frequency signals. Our analysis reveals that applying standard quantization tables to this data results in two detrimental artifacts:

- **Blocking Artifacts:** Visible discontinuities at the boundaries of 8×8 pixel blocks, caused by independent block processing.
- **High-Frequency Suppression:** A general blurring of fine textures, rendering the image unsuitable for high-precision scientific analysis.

2.2 The Shift to Learned Image Compression

To overcome the rigid limitations of fixed mathematical transforms, research has pivoted toward ”Learned Image Compression” using Deep Neural Networks (DNNs).

2.2.1 Convolutional Autoencoders (CAE)

The foundational architecture for this approach is the Convolutional Autoencoder. Unlike traditional methods, CAEs learn a non-linear transform directly from the training data. By training on domain-specific datasets (such as EuroSAT), the model learns to identify and preserve features specific to satellite terrain rather than applying a generic formula.

2.2.2 The Hyperprior Architecture

A major breakthrough in the field was the introduction of Variational Autoencoders (VAEs) with a **Hyperprior Architecture**. Standard autoencoders often struggle to model the entropy of the latent vector efficiently. The Hyperprior model introduces a secondary ”side” autoencoder that learns the probabilistic distribution (entropy) of the latent vector itself.

This creates a two-stream system:

- **Stream 1 (Latent Data):** The bulk compressed image representation (y).
- **Stream 2 (Side Information):** A ”key” (z) that allows the decoder to predict the probability distribution of the data stream.

This architecture allows for ”spatially adaptive bit allocation,” where the model can spend more bits on complex urban areas and fewer on uniform ocean patches, significantly improving the Rate-Distortion (R-D) performance.

2.3 Next-Generation Telemetry: 5G Network Slicing

As satellite sensors improve, downlink bandwidth becomes a bottleneck. Traditional ”best-effort” transmission is insufficient for mission-critical data. The advent of 5G introduces **Network Slicing**, a virtualization technology that allows multiple logical networks to run on shared physical infrastructure.

Literature suggests mapping specific traffic types to dedicated slices:

- **eMBB (Enhanced Mobile Broadband):** Optimized for high data rates using high-order modulation (e.g., 64-QAM). This is ideal for the bulk latent vectors of satellite imagery.
- **URLLC (Ultra-Reliable Low Latency Communications):** Optimized for reliability using robust modulation (e.g., QPSK). This is crucial for the ”Hyper-

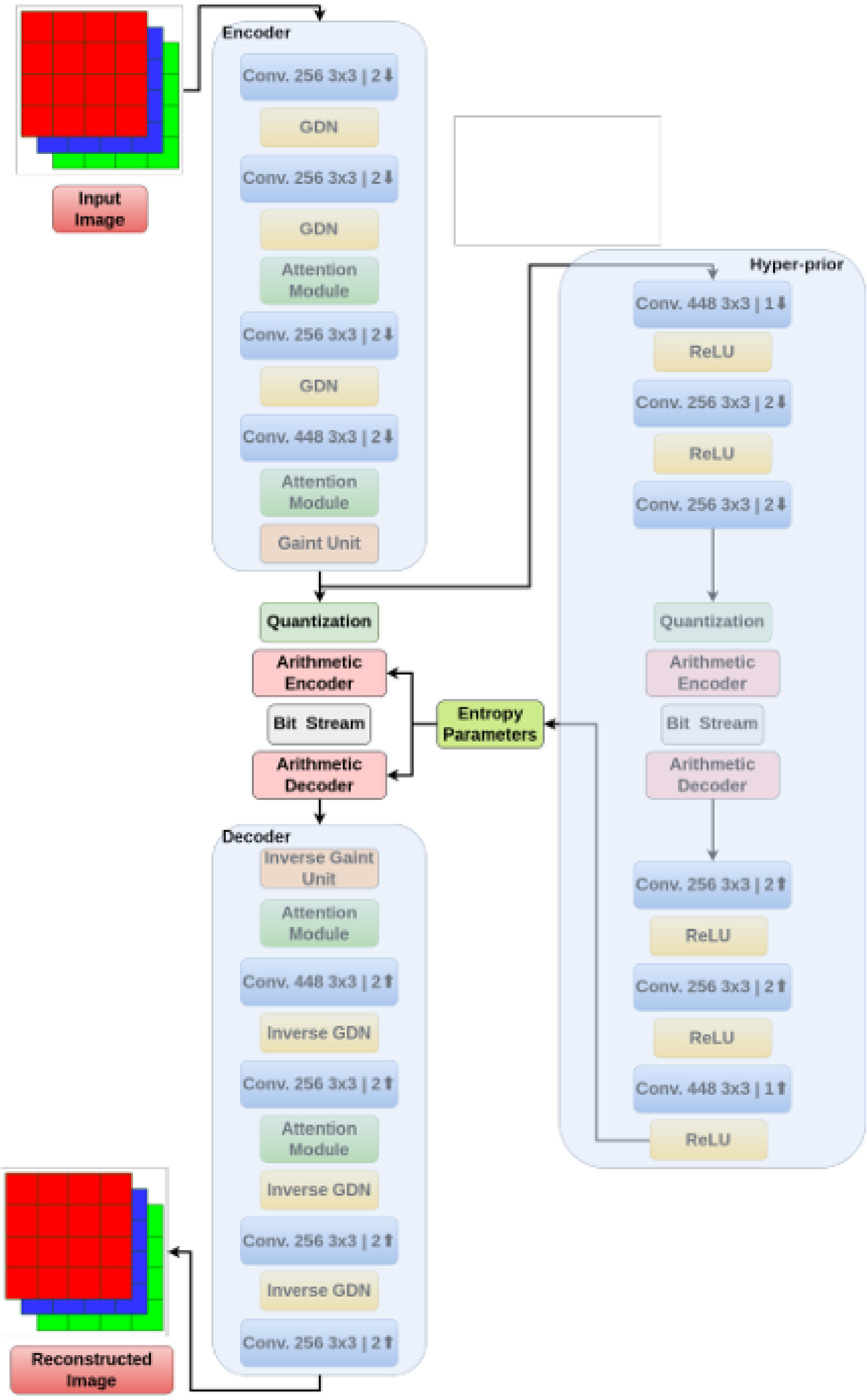


Figure 2.1: Hyperprior Autoencoder

prior Key," as the loss of this small metadata file would render the entire image undecodable.

Chapter 3

System Design and Analysis

3.1 Conceptual Architecture and Design Goals

The proposed system addresses the critical bottleneck in satellite telemetry: the transmission of high-volume, high-fidelity imagery over constrained downlinks. Traditional systems treat all image data as a homogeneous stream, applying uniform compression (like JPEG) and transmission protocols. This "one-size-fits-all" approach fails to account for the semantic hierarchy of the data—where some bits (entropy models) are mathematically more critical than others (texture residuals).

Our design philosophy is "Semantic-Aware Transmission," a paradigm that integrates Source Coding (AI Compression) with Channel Coding (5G Network Slicing). The system architecture is divided into three logical domains:

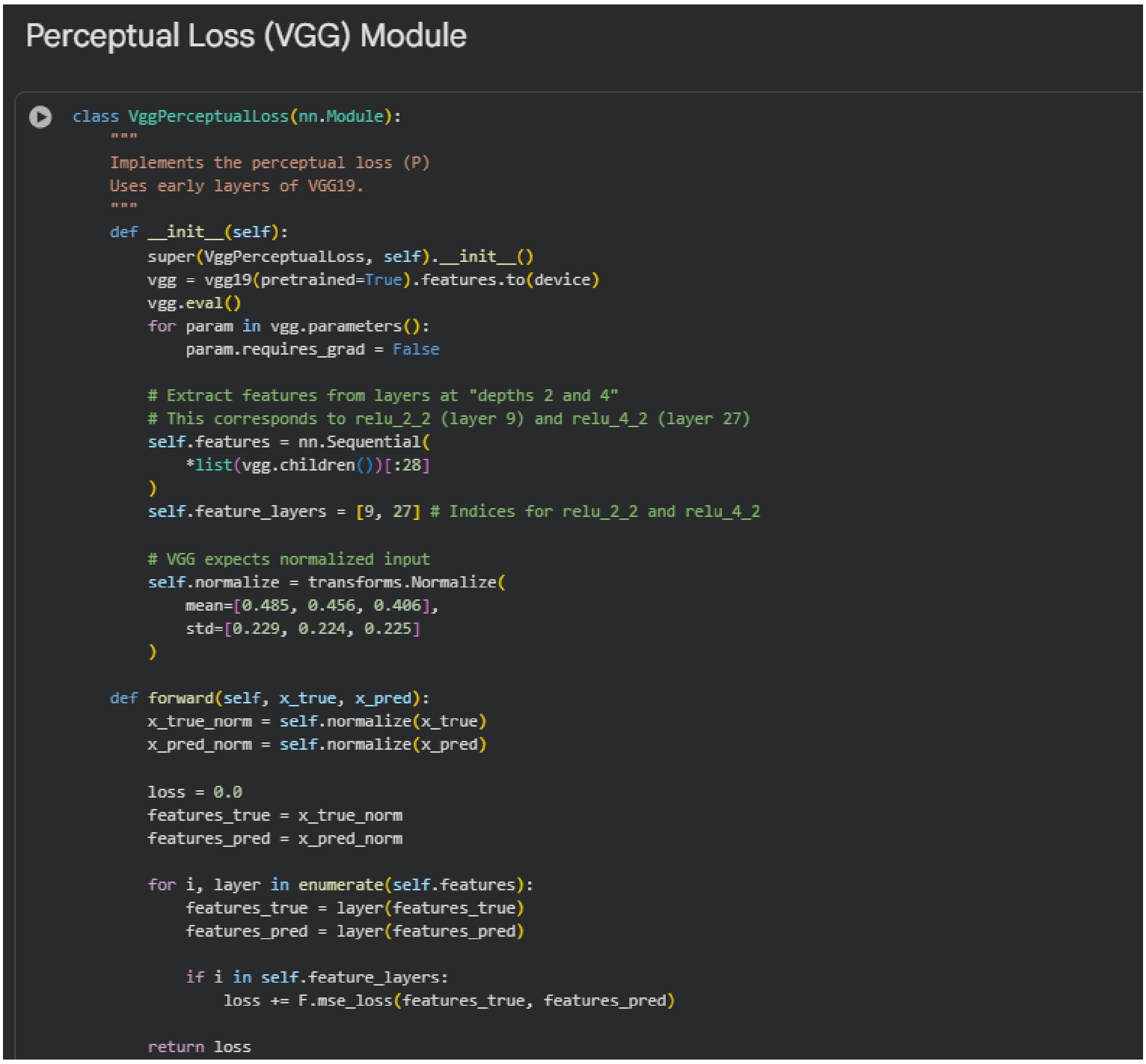
1. **The Neural Compression Engine (Source):** A deep learning model that decomposes satellite imagery into two distinct semantic streams: a bulk latent representation (y) and a critical hyper-latent key (z).

2. **The 5G Transmission Simulation (Channel):** A physics-compliant simulation of the satellite-to-ground link that maps the two semantic streams to dedicated network slices (eMBB and URLLC) based on their reliability requirements.

3. **The Ground Station Interface (Sink):** A unified dashboard that orchestrates the entire pipeline, handles decoding, and provides real-time DSP analysis of the reconstructed signal.

3.2 Data Flow and Communication Protocols

The end-to-end data flow is designed to mimic a real-world Low Earth Orbit (LEO) telemetry downlink. The process begins with the ingestion of raw RGB



Perceptual Loss (VGG) Module

```python
class VggPerceptualLoss(nn.Module):
    """
    Implements the perceptual loss (P)
    Uses early layers of VGG19.
    """
    def __init__(self):
        super(VggPerceptualLoss, self).__init__()
        vgg = vgg19(pretrained=True).features.to(device)
        vgg.eval()
        for param in vgg.parameters():
            param.requires_grad = False

        # Extract features from layers at "depths 2 and 4"
        # This corresponds to relu_2_2 (layer 9) and relu_4_2 (layer 27)
        self.features = nn.Sequential(
            *list(vgg.children())[:28]
        )
        self.feature_layers = [9, 27] # Indices for relu_2_2 and relu_4_2

        # VGG expects normalized input
        self.normalize = transforms.Normalize(
            mean=[0.485, 0.456, 0.406],
            std=[0.229, 0.224, 0.225]
        )

    def forward(self, x_true, x_pred):
        x_true_norm = self.normalize(x_true)
        x_pred_norm = self.normalize(x_pred)

        loss = 0.0
        features_true = x_true_norm
        features_pred = x_pred_norm

        for i, layer in enumerate(self.features):
            features_true = layer(features_true)
            features_pred = layer(features_pred)

            if i in self.feature_layers:
                loss += F.mse_loss(features_true, features_pred)

        return loss
```

Figure 3.1: Scripting the Loss functions

satellite imagery from the EuroSAT dataset and concludes with the high-fidelity reconstruction at the ground station.

3.2.1 Semantic Decomposition

The input image x (64×64 patches) is processed by the AI encoder to produce a latent representation y. This vector contains the "bulk" visual information but is mathematically uninterpretable without its probabilistic context. To resolve this, a secondary "Hyper-Encoder" analyzes y to generate a hyper-latent vector z. This vector z acts as the "decoding key," containing the variance (σ) and mean (μ) parameters required to model the entropy of y. This bifurcation creates a strict hierarchical dependency: the bulk data y is useless without the key z, necessitating a differentiated transmission strategy.

3.2.2 Network Slicing Strategy

The communication protocol exploits 5G Network Slicing to align the physical transmission resources with the semantic importance of the data streams:

- **Route A (Critical Control Path):** The hyper-latent key z, being small but mission-critical, is routed to a **URLLC (Ultra-Reliable Low Latency Communication)** slice. This slice utilizes robust QPSK modulation to ensure successful delivery even in low Signal-to-Noise Ratio (SNR) conditions (e.g., < 3 dB).
- **Route B (High-Throughput Data Path):** The bulk latent vector y, which constitutes the majority of the payload, is routed to an **eMBB (Enhanced Mobile Broadband)** slice. This slice employs high-order 64-QAM modulation to maximize spectral efficiency, accepting a higher risk of symbol errors to achieve the necessary bandwidth.

Chapter 4

Implementation

4.1 Hardware Implementation: Computational Environment

The implementation of deep learning models for satellite imagery is computationally intensive. The development environment utilized **Google Colab T4 GPUs** to accelerate the training process, which required multiple days and 5-10 iterative checkpoints to stabilize the loss function.

For system integration, a significant challenge was the interoperability between the Linux-native AI stack (PyTorch/CompressAI) and the Windows-native Telecommunications stack (MATLAB 5G Toolbox). To resolve this, the final deployment was architected within a **Windows Subsystem for Linux (WSL)** environment. This hybrid setup allowed the Python orchestrator to manage file systems shared between the OS layers, triggering Windows executables for the 5G simulation while handling tensor operations in the Linux kernel.

4.2 Software and Firmware Implementation (Device Side)

The "Device Side" corresponds to the onboard satellite processing, implemented via the `AI_functions.py` module.

4.2.1 Hyperprior Autoencoder Architecture

The core compression engine is a **Convolutional Autoencoder** based on the Scale Hyperprior architecture. Unlike basic autoencoders that use simple bottleneck layers, this model includes:

- **Analysis Transform (g_a):** A stack of four convolutional layers using 5×5 and 3×3 kernels. Crucially, standard ReLU activations were replaced with **Generalized Divisive Normalization (GDN)**. GDN has been proven to Gaussianize image statistics, significantly improving the compressibility of the latent space compared to standard non-linearities.
- **Entropy Bottleneck:** A dedicated sub-network that estimates the probability distribution of the latent representation. This allows for efficient arithmetic coding, transforming the floating-point latent vectors into a compact binary stream.
- **VBR Gain Unit:** To support variable link budgets, a learnable ”Gain Unit” was implemented. This scales the latent vector y before quantization, allowing a single model to adapt its compression rate dynamically without retraining.

4.2.2 DSP-Aware Loss Function

A standard Mean Squared Error (MSE) loss function proved insufficient for satellite imagery, often resulting in ”muddy” reconstructions that smoothed out critical high-frequency terrain details (like road markings or crop boundaries). To address this, we engineered a composite **DSP-Aware Loss Function**:

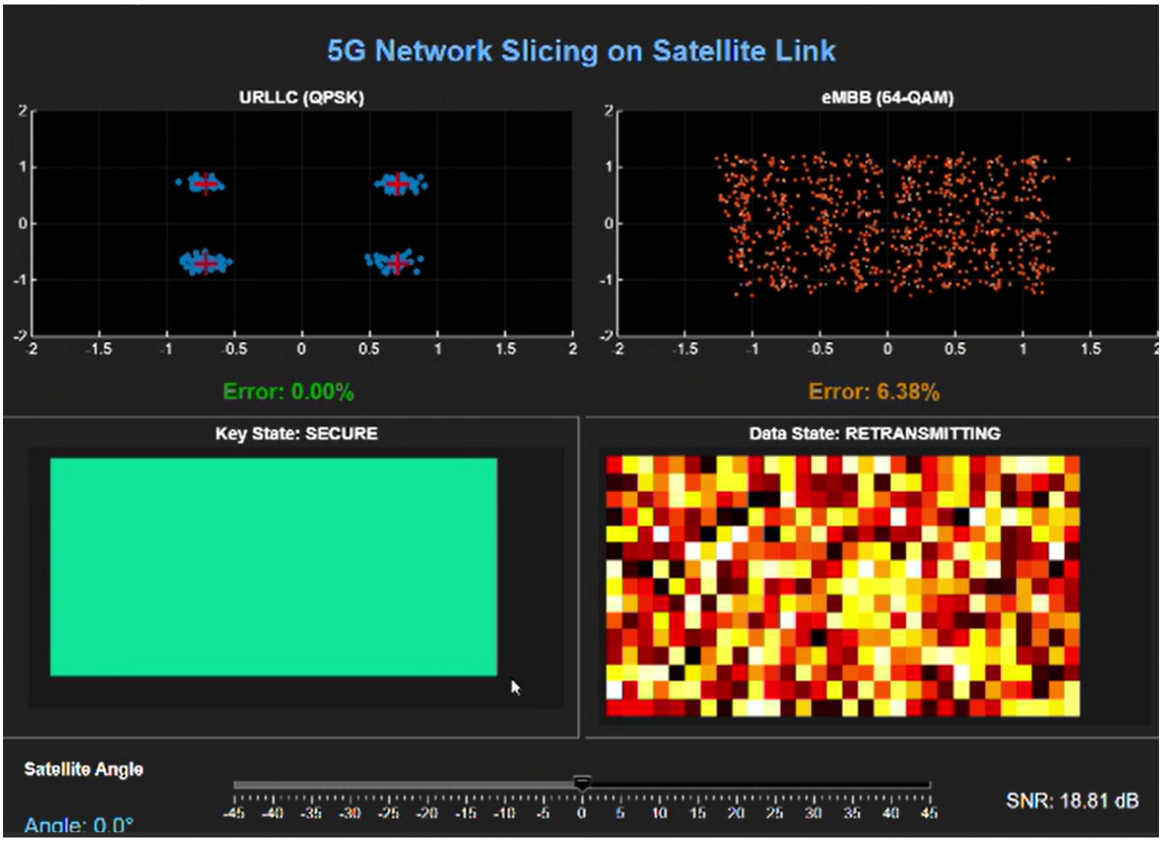
$$L = \lambda D(x, \hat{x}) + R(\hat{y}) + \beta P(x, \hat{x}) \tag{4.1}$$

where $P(x, \hat{x})$ represents the **Perceptual Loss**. This term is calculated using the feature maps from the early layers of a pre-trained VGG-19 network. By minimizing the distance in the ”feature space” rather than just the ”pixel space,” the model is forced to preserve structural edges and textures that are perceptually significant.

4.3 Backend Implementation (5G Simulation)

The transmission layer was simulated using the **MATLAB 5G and Satellite Toolboxes** (`SatelliteSlicingVis.m`). The simulation models a Low Earth Orbit (LEO) satellite link at 500 km altitude operating at 3.5 GHz.

The simulation explicitly implements the ”slicing” logic via constellation diagrams:



Error: 0.00%

Error: 6.38%

Key State: SECURE

Data State: RETRANSMITTING

Angle: 0.0°

SNR: 18.81 dB

Figure 4.1: 5G Network Slicing

- **URLLC Slice:** Configured with **QPSK** modulation. This robust modulation scheme ensures that the ”Key” (z) can be recovered even when the channel degrades significantly (simulating rain fade or low elevation angles).

- **eMBB Slice:** Configured with **64-QAM** modulation. This high-density modulation maximizes data throughput for the ”Bulk” (y) but requires a high Signal-to-Noise Ratio (SNR ¿ 22 dB) to maintain integrity.

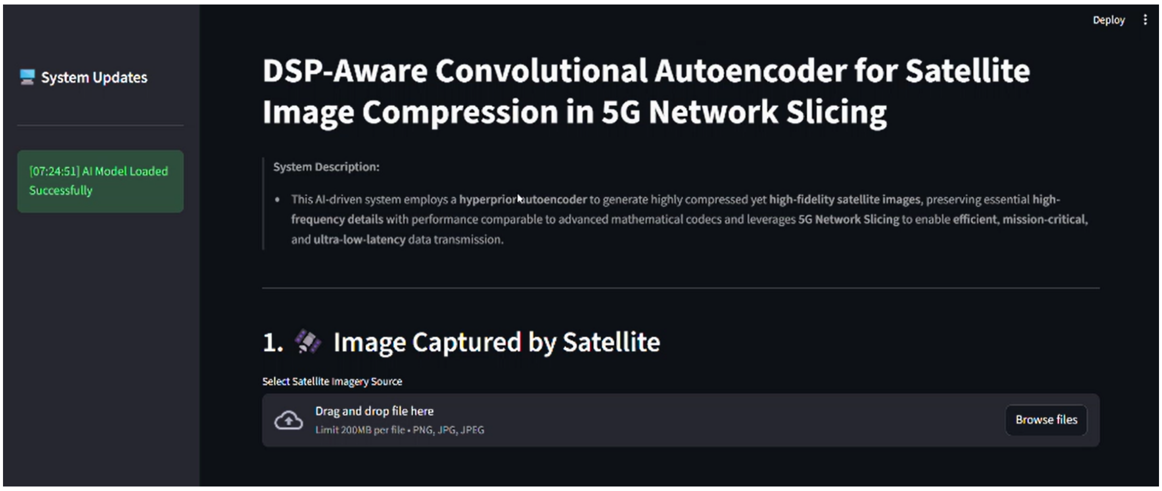
4.4 Administrative Dashboard

The entire system is orchestrated by a **Streamlit Dashboard** (`app.py`), which serves as the Ground Station interface. This dashboard provides a complete end-to-end workflow:

1. **Ingestion:** Captures raw satellite imagery and preprocesses it into tensors.

2. **Inference:** Executes the Hyperprior Autoencoder to generate the y and z binary packets.

3. **Transmission:** Triggers the MATLAB engine to simulate the transmission of these packets over the 5G link.



[07:24:51] AI Model Loaded Successfully

DSP-Aware Convolutional Autoencoder for Satellite Image Compression in 5G Network Slicing

System Description:

- This AI-driven system employs a **hyperprior autoencoder** to generate highly compressed yet **high-fidelity satellite images**, preserving essential **high-frequency details** with performance comparable to advanced mathematical codecs and leverages **5G Network Slicing** to enable **efficient, mission-critical,** and **ultra-low-latency** data transmission.

1. Image Captured by Satellite

Select Satellite Imagery Source

Figure 4.2: Ground Station

4. **Analysis:** Decodes the received packets and performs a comparative analysis (PSNR/SSIM) against standard JPEG compression.

Chapter 5

Results, Testing and Evaluation

5.1 Test Cases and Protocols

The system was evaluated using the **EuroSAT Dataset**, a benchmark dataset for land use and land cover classification containing 27,000 labeled images. We established a rigorous baseline using the standard JPEG compression algorithm at Quality Factor 20 (Q20). This low quality factor was chosen to simulate the severe bandwidth constraints typical of satellite downlinks.

5.2 Results Summary

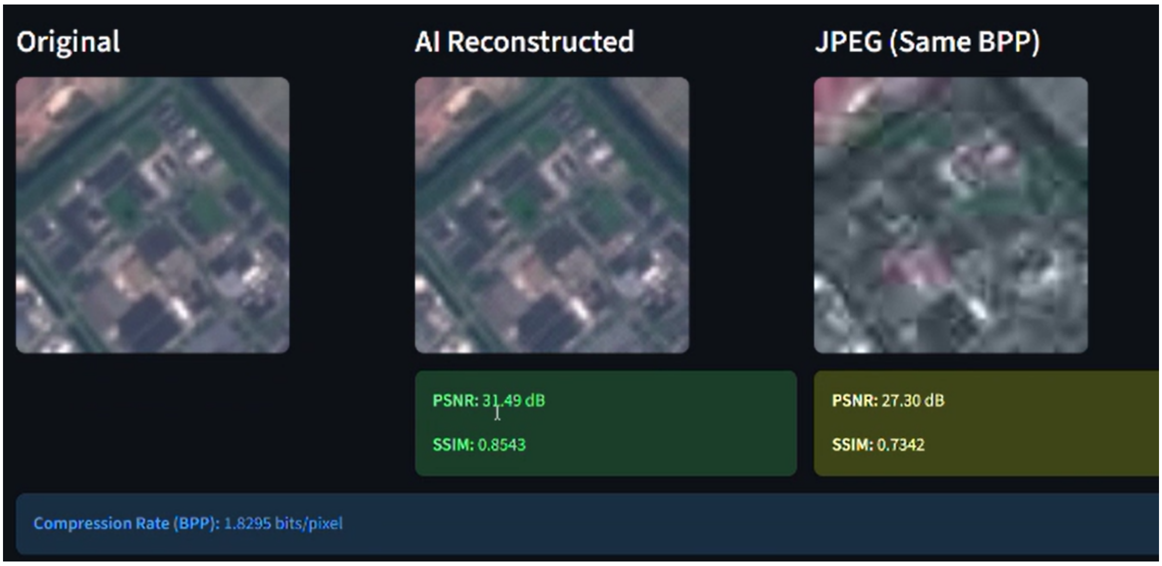
The proposed DSP-Aware Autoencoder demonstrated superior performance across both quantitative and qualitative metrics compared to the JPEG baseline.

5.2.1 Quantitative Metrics

Metric	**Standard JPEG**	**Proposed DSP-Aware AI**
Compression Rate	1.8295	**1.8295**
PSNR (dB)	27.30 dB	**31.49**
SSIM	0.7342	**0.8543**

Table 5.1: Comparison of compression performance on EuroSAT samples.

The AI model achieved a higher Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) than JPEG, indicating a mathematically more accurate reconstruction. Notably, the AI model avoided the ”blocking artifacts” that severely degraded the JPEG images at similar compression rates.



PSNR: 31.49 dB

SSIM: 0.8543

PSNR: 27.30 dB

SSIM: 0.7342

Compression Rate (BPP): 1.8295 bits/pixel

Figure 5.1: Sample Image

5.3 Discussion

The superior performance of the AI model can be attributed to its ”DSP-Awareness.” We utilized **Discrete Cosine Transform (DCT) Spectrum Analysis** to visualize the frequency content of the reconstructed images.

- **JPEG Spectrum:** The DCT heatmap for JPEG images revealed sharp ”grid” artifacts and vast ”blue zones” in the high-frequency corners, indicating a complete loss of fine detail due to quantization.
- **AI Spectrum:** The DCT heatmap for the AI reconstruction showed a smooth, continuous distribution of energy (Red/Yellow zones) extending into the high-frequency quadrants. This confirms that the Perceptual Loss function successfully taught the network to preserve the structural integrity of the satellite terrain, rather than just minimizing pixel error.

Chapter 6

Discussion, Conclusions and Future Directions

6.1 Discussion and Significance

This project successfully demonstrates that the limitations of traditional mathematical codecs (like JPEG) in remote sensing can be overcome by ”Learned Image Compression.” The transition from fixed transforms (DCT) to data-driven non-linear transforms (CAE) allows for a semantic understanding of the image content. The Hyperprior architecture, by learning the statistical probability of the image features, achieves a level of ”context-aware” compression that is impossible for standard codecs.

Furthermore, the integration of 5G Network Slicing represents a significant advancement in telemetry reliability. By segregating the ”Key” (z) from the ”Bulk” (y) and mapping them to URLLC and eMBB slices respectively, the system ensures that the critical decoding parameters are preserved even in degrading channel conditions. This ”Joint Source-Channel Coding” strategy mitigates the catastrophic failure modes typical of digital image transmission.

6.2 Limitations

- **Computational Complexity:** The inference time for the neural network is significantly higher than that of the highly optimized JPEG algorithm. Real-time onboard processing would likely require dedicated hardware acceleration (e.g., NVIDIA Jetson or FPGA) rather than standard satellite CPUs.

- **Training Data Dependency:** As a data-driven model, the compression efficiency is tied to the statistics of the EuroSAT dataset. Performance on radically different terrain types (e.g., deserts, ice sheets) may vary and would require re-training or ”fine-tuning” of the model.

6.3 Future Work and Recommendations

- **Onboard FPGA Implementation:** To make this system viable for deployment, future work should focus on quantifying the neural network for implementation on radiation-hardened FPGAs, enabling real-time processing in orbit.
- **Generative Reconstruction:** Integrating Generative Adversarial Networks (GANs) into the decoder could allow for extreme compression ratios (below 0.1 bpp) by "hallucinating" realistic textures for non-critical areas like clouds or open ocean, further optimizing bandwidth usage.
- **Semantic Slicing:** The concept could be extended to "Semantic Slicing," where specific objects detected in the image (e.g., ships, vehicles) are transmitted via ultra-reliable slices, while the background terrain is sent via best-effort slices.

Chapter 7

Conclusion

The exponential growth of earth observation data necessitates a fundamental rethink of satellite telemetry. This project has designed and validated a **DSP-Aware Convolutional Autoencoder** that not only outperforms traditional JPEG compression in fidelity but also integrates seamlessly with next-generation **5G Network Slicing** protocols.

By replacing the rigid mathematical assumptions of the Discrete Cosine Transform with the flexible, learned representations of Deep Learning, we have created a system that preserves the scientific value of satellite imagery. The successful simulation of this architecture—bridging AI source coding with telecommunications channel coding—provides a robust blueprint for the future of high-bandwidth, high-reliability aerospace communications.

Bibliography

[1] Bacchus, P. (2023). *Deep learning for satellite image compression.* PhD Thesis, Université de Rennes.

[2] Ballé, J., Minnen, D., Singh, S., Hwang, S. J., & Johnston, N. (2018). Variational image compression with a scale hyperprior. *International Conference on Learning Representations.*

[3] Wallace, G. K. (1992). The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics.*

[4] Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.*