Deep Learning-Based Compressed Sensing for Drone Image Transmission

Abdelmouhaimen Sarhane
INP-ENSEEIHT

Abstract

This work explores an innovative approach to image compression and reconstruction using compressed sensing and deep learning. The objective is to develop an adaptive sampling strategy for drone images, leveraging deep learning models for both mask generation and image inpainting. We introduce an adaptive binarization method using mean shift, incorporating a contour-weighting strategy controlled by a lambda parameter. The system enables real-time adaptive sampling and reconstruction, enhancing efficiency in bandwidth-limited environments such as drone-to-ground station transmissions. We validate our approach with a proof-of-concept implementation using the Streamlit framework.

Keywords: Compressed Sensing, Deep Learning, Image Reconstruction, Mask Generation, Streamlit.

1 Introduction

Drones are increasingly used in surveillance, mapping, and environmental monitoring, but their ability to transmit high-resolution images in real time is constrained by bandwidth limitations. Traditional image compression methods, such as JPEG or HEVC, reduce file size but require full image acquisition before compression. Compressed Sensing (CS) is a promising alternative, allowing images to be sampled at rates below the Nyquist limit while enabling accurate reconstruction.

In this work, we explore a deep learning-based CS approach to adaptively sample images from a drone while ensuring effective reconstruction at the ground station. Unlike static sampling methods, our approach uses a convolutional neural network (CNN) to generate an optimized sampling mask, dynamically adjusting based on scene content. Inspired by [2], we introduce a contour-aware mask generation strategy to preserve essential features in low-sampling regimes.

2 Model and Mask Generation

2.1 Neural Networks for Image Processing

We employ two models: *NetE* for image inpainting and *NetM* for adaptive mask generation.

- **NetE**: U-Net inspired autoencoder trained to reconstruct missing pixels.
- **NetM**: Residual CNN trained to generate binary masks at different sampling rates.

NetE reconstructs images by learning to inpaint missing regions using convolutional layers, batch normalization, and non-linear activations. NetM, on the other hand, generates an optimized sampling mask that maximizes image reconstruction accuracy while maintaining sparsity constraints.

2.2 Architecture

2.2.1 NetE: Image Reconstruction Model

NetE is an autoencoder-based neural network designed to inpaint missing pixels from sampled images. Inspired by U-Net, it consists of an encoder-decoder structure with convolutional layers, batch normalization, and skip connections to preserve spatial features. Figure 1 illustrates its architecture.

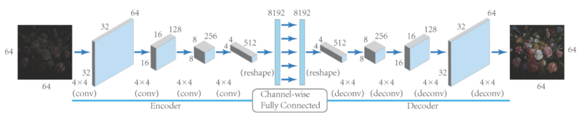

Figure 1: NetE: Autoencoder-based reconstruction network.

2.2.2 NetM: Adaptive Mask Generation Model

NetM generates binary sampling masks using a deep residual CNN. The network learns to optimize pixel selection for compressed sensing using supervised learning. The final mask is binarized using mean shift to ensure a controlled sampling rate. Figure 2 presents its architecture.

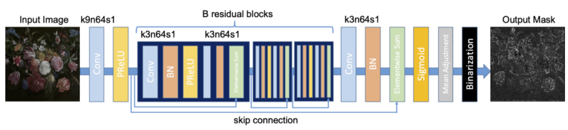

Figure 2: NetM: Residual CNN for adaptive mask generation.

2.3 Mean Shift Binarization

To enforce sparsity constraints, we use *mean shift* binarization to control sampling density:

$$B = \mathrm{Ber}(\frac{c}{\overline{L}}L), \tag{1}$$

where B is the binary mask, L is the mask prediction, and c is the target sparsity. The mean shift method ensures that the generated masks maintain a balance between preserving important structures and reducing redundancy.

2.4 Contour Enhancement with Lambda Weighting

To prioritize edge preservation, a lambda constant λ adjusts the weight of contours in the sampling mask:

$$B' = B + \lambda \nabla I, \tag{2}$$

where ∇I is the image gradient. This method has been successfully used in scientific imaging applications such as X-ray fluorescence image reconstruction [2], where sharp transitions are critical for preserving structural integrity.

3 Mask Update and Transmission Strategy

3.1 Mask Transmission Process

At the start, the ground station sends an initial random mask to the drone. This mask is used for sampling transmitted images. Over time, the mask is progressively refined based on received images.

4 Mask Update and Transmission Strategy

4.1 Mask Transmission Process

Initially, the ground station transmits a random mask to the drone, which uses it for image sampling. The mask is progressively refined based on received images to improve reconstruction quality.

4.2 Periodic Original Image Transmission for Mask Refinement

To ensure adaptability, the drone periodically transmits an uncompressed image. The ground station accumulates multiple images to update the mask dynamically.

Steps:

1. Ground station reconstructs received compressed images.
2. Periodically, an original uncompressed image is transmitted.
3. Multiple images are used to compute an updated mask.
4. A new optimized mask is sent back to the drone for improved sampling.

This iterative process ensures an adaptive sampling strategy while minimizing bandwidth usage.

5 Implementation and Practical Considerations

5.1 Image Sampling and Transmission

Each frame is transmitted using the following format:

$$\text{ByteCode} + \text{len}(\text{data}) + (\text{Nb}_{\text{pixels}}, 3). \tag{3}$$

The station updates masks every N frames based on an accumulated history of past images.

5.2 Latency and Mask Synchronization

Since the station sends new masks asynchronously, there exists a delay in mask propagation. To address this, the drone appends a *mask ID* to sampled images, ensuring proper synchronization between mask usage and reconstruction.

5.3 GUI and Proof-of-Concept

A real-time prototype was implemented using Streamlit to visualize sampling, mask generation, and reconstruction. Figure 3 presents a screenshot of the interface.

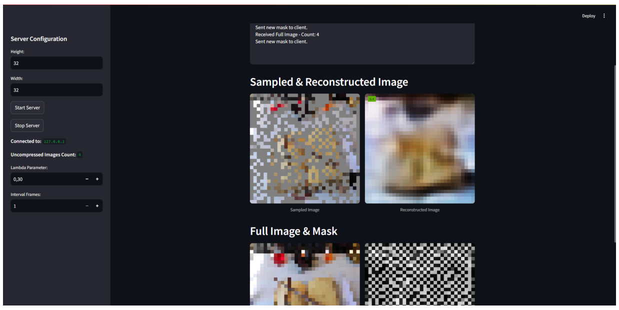

Figure 3: Streamlit-based GUI for real-time visualization.

6 Results and Discussion

6.1 Performance at Different Sampling Rates

We evaluated NetE and NetM at sampling rates of 20% and 50%. Figure 4 presents reconstruction quality at these settings.

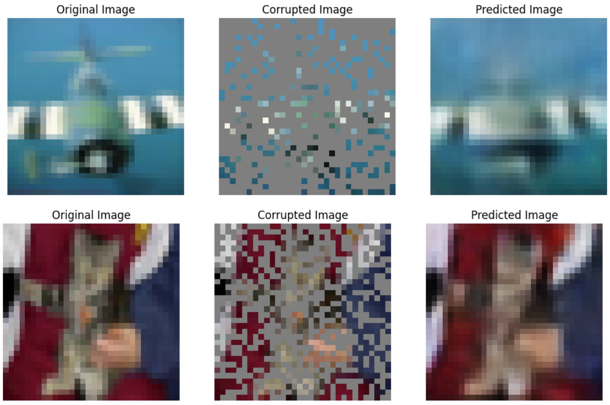

Figure 4: Reconstruction quality comparison: 20% (left) vs. 50% (right).

Figure 5: Mask generated by NetM at 50% sampling rate.

6.2 Training Performance and Convergence

The loss curves in Figure 6 indicate stable convergence, demonstrating effective model training.

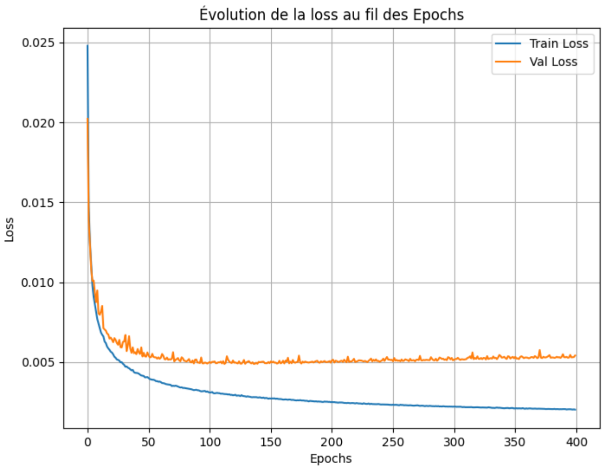

Figure 6: Loss evolution over training epochs.

6.3 Challenges and Optimizations

- **Low-resolution constraints**: Due to computational limits, 32x32 images were used, which affects fine details.
- **Latency in mask updates**: Implementing buffer-based synchronization improved reliability.
- **Edge loss in sparse masks**: Contour weighting significantly improved feature retention.

7 Conclusion and Future Work

This paper introduced a deep learning-based compressed sensing approach for drone image transmission. The combination of NetM-generated masks and NetE-based reconstruction showed significant improvements over random sampling. Our periodic mask update strategy ensured long-term efficiency.

Future improvements include:

- Scaling to high-resolution images.
- Implementing real-time embedded processing on drones.
- Exploring reinforcement learning for adaptive sampling.

These advancements will enhance real-world applicability in drone imaging.

Acknowledgements

We thank our advisors and colleagues for their guidance in developing this system.

References

[1] Machidon, A.L., Pejović, V. (2023). "Deep learning for compressive sensing: a ubiquitous systems perspective." Artificial Intelligence Review, 56, 3619–3658.

[2] "Adaptive Image Sampling using Deep Learning and its Application on X-Ray Fluorescence Image Reconstruction." (Year). Journal/Conference Name.