# Weekly Update: Drone Image Evaluation Results

## 1. Summary of Results (PHASE 1)

Original Hypothesis was: There will be less data loss compared to UDP. However based on results, it seems on good networks reconstruction performs worse. But on moderate to high congestion. The model performs better than baseline transmission.

The evaluation was run on custom dataset (~240 images), which contains images from drones.
The only time reconstruction is better than what is recieved via UDP when when a packet loss is greater than 20% or original packets sent.

PSNR and SSIM for flowmo seem to be lower than all other cases (which is bad). Thus supporting our hypothesis.

Overall there is more loss when reconstructing the image. But the semantic loss is much lower as seen in LPIPS.

# 2. Understanding LPIPS (Learned Perceptual Image Patch Similarity)

LPIPS, or Learned Perceptual Image Patch Similarity, is an image quality metric designed to align more closely with human perception of image similarity than traditional metrics like Peak Signal-to-Noise Ratio (PSNR) or Structural Similarity Index (SSIM).

Instead of comparing pixel values directly, LPIPS leverages deep learning. It computes the difference between features extracted from two images using pre-trained convolutional neural networks (like AlexNet or VGG). These features, learned by the networks for tasks like image classification, tend to capture more abstract and perceptually relevant aspects of an image. A lower LPIPS score indicates that the two images are more perceptually similar.

# 3. LPIPS Feature Differences: AlexNet vs. VGG

When LPIPS is used, the choice of the underlying pre-trained network (the "backbone") affects the types of features being compared:

- **LPIPS with AlexNet (lpips_alex)**:
  - AlexNet is a relatively shallower convolutional neural network compared to VGG.
  - LPIPS using AlexNet is generally faster to compute.
  - It tends to capture more global, texture-like, and lower-level features. Differences in these aspects between images will be more heavily penalized.
- **LPIPS with VGG (lpips_vgg)**:
  - VGG is a deeper network architecture.
  - LPIPS using VGG is often considered a more robust perceptual metric as deeper networks can capture more complex and semantic information, as well as finer details.

- It may be more sensitive to structural differences and object-level distortions.

By examining both AlexNet-based and VGG-based LPIPS scores, we can get a more comprehensive understanding of the perceptual differences. For instance, one model might perform well on texture (AlexNet) but poorly on finer structural details (VGG), or vice-versa.

# 4. Decoding the Evaluation Conditions

The evaluation involved several conditions to test the model:

- **`flow_mo`**: This condition (renamed from `model_reconstruction` in the plots) represents the baseline performance of the FlowMo model. It shows the quality of images reconstructed by the model directly, without any simulated UDP packet loss.
- **`udp_lrX_psY`**: These conditions simulate packet loss over a UDP.
    - `lrX`: Denotes the "loss rate." For example, `lr0.05` means a 0.5% packet loss rate (i.e., 5% of the simulated packets are dropped).`lr0.4` means 40% and so on.
    - `psY`: Indicates the "patch size." For example, `ps8` means the image was conceptually divided into 8x8 pixel patches, and loss was simulated by dropping entire patches. Other patch sizes used were `ps16` (16x16), `ps32` (32x32), and `ps64` (64x64).

- **Relation to Packet Size**: The `patch_size` directly relates to the simulated packet size. A larger `patch_size` (e.g., `ps64`) means that if a "packet" (patch) is lost, a larger contiguous block of the image is gone. This simulates the scenario where larger data packets are dropped during transmission. Conversely, a smaller `patch_size` (e.g., `ps8`) simulates the loss of smaller data packets.

# Comparison of packet loss with real world condition

| Environment | Typical UDP Packet Loss % | Notes |
|---|---|---|
| **Less noisy (Good Wi-Fi / Ethernet)** | **0–1%** | Usually negligible; minor loss may occur due to congestion or buffer overflow. |
| **Moderate noise (Public Wi-Fi / Busy LAN)** | **1–5%** | Occasional jitter, latency, and dropped packets — acceptable for streaming, but not critical data. |

| Environment | Typical UDP Packet Loss % | Notes |
| --- | --- | --- |
| **Highly noisy (Mobile networks, saturated Wi-Fi, long-distance wireless)** | **5–30%+** | Packet loss can spike; real-time apps (VoIP, gaming) can break or degrade. |

# 5. Visual Results

The following plots illustrate the aggregated performance metrics across the different conditions.

*Note: Lower LPIPS scores are better. Higher PSNR and SSIM scores are better.*

## LPIPS Alex Comparison

LPIPS Alex Comparison

## LPIPS VGG Comparison

LPIPS VGG Comparison

## PSNR Comparison

PSNR Comparison

## SSIM Comparison

SSIM Comparison

# For Phase 2:

We can either:

- Work on improving the model by finetuning and making it more noise resistant.
- Make it perform better on text regeration.
- Optimise the diffusion process and make it more efficient.

## (Need to discuss on what direction to move further.)