

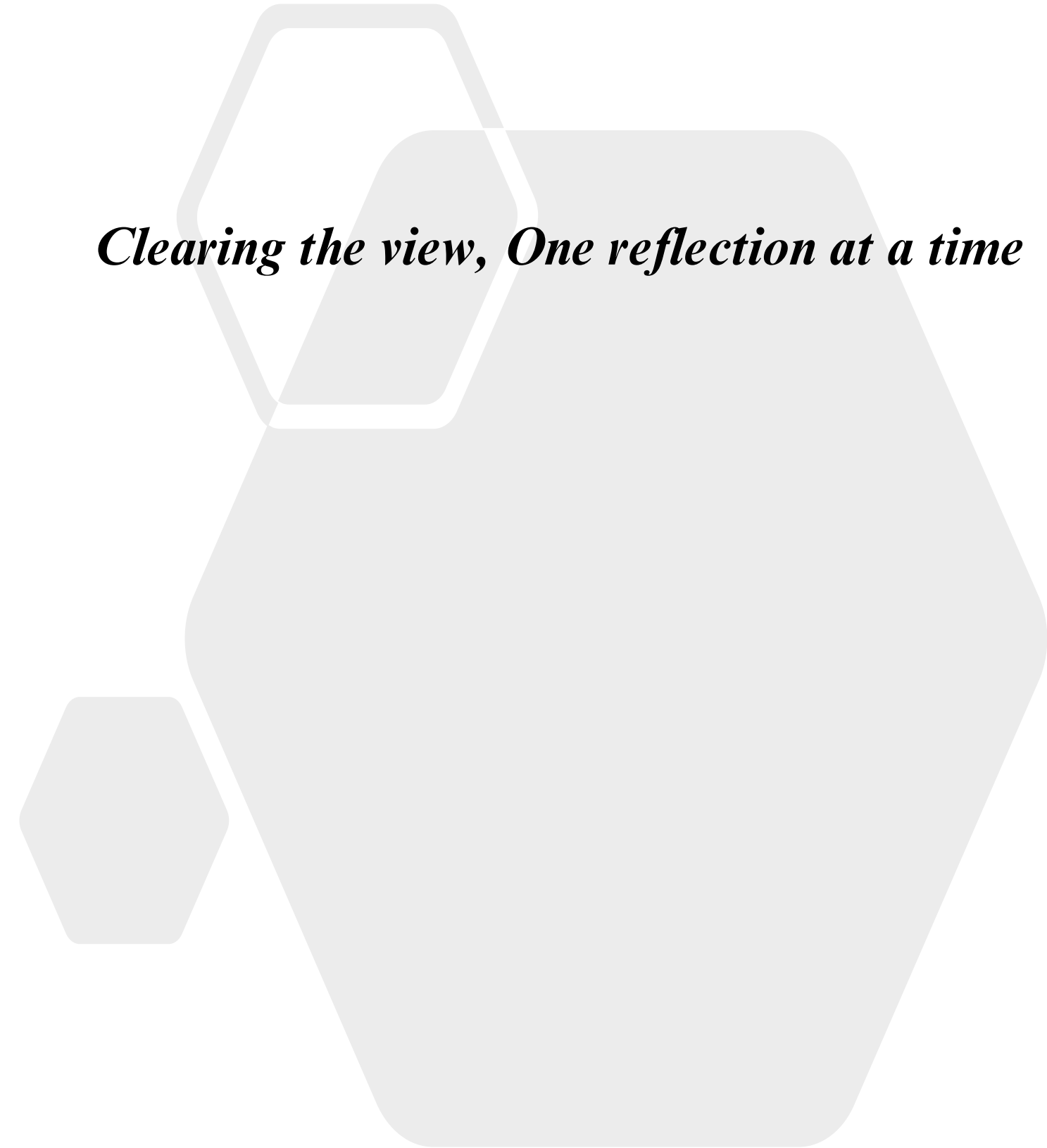
# QUALCOMM VISIONX

## Aether

*Clearing the view, One reflection at a time*

**Problem Statement** - Development innovative AI models for reflection removal, targeting real-world applications.

**Team Name** - Aether

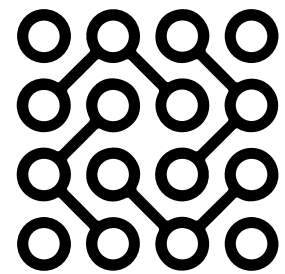


# Aether

*Clearing the view, One reflection at a time*

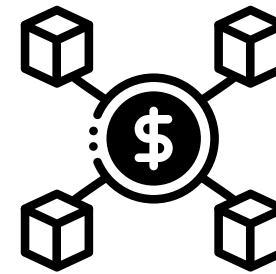


Aether is a **state-of-the-art deep learning enabled** platform for removing reflections. Unlike traditional solutions to the problem, we are utilizing **advanced VQGAN with LORA modules** and fine-tuning the encoder for reflection removal



## Ensuring generalization

Finetuning **encoders** with **LORA modules** ensures zero-shot deployment on any input image



## Maintaining Quality

Our solution removes reflections while **preserving the integrity** of the original content, using a frozen dictionary of embeddings



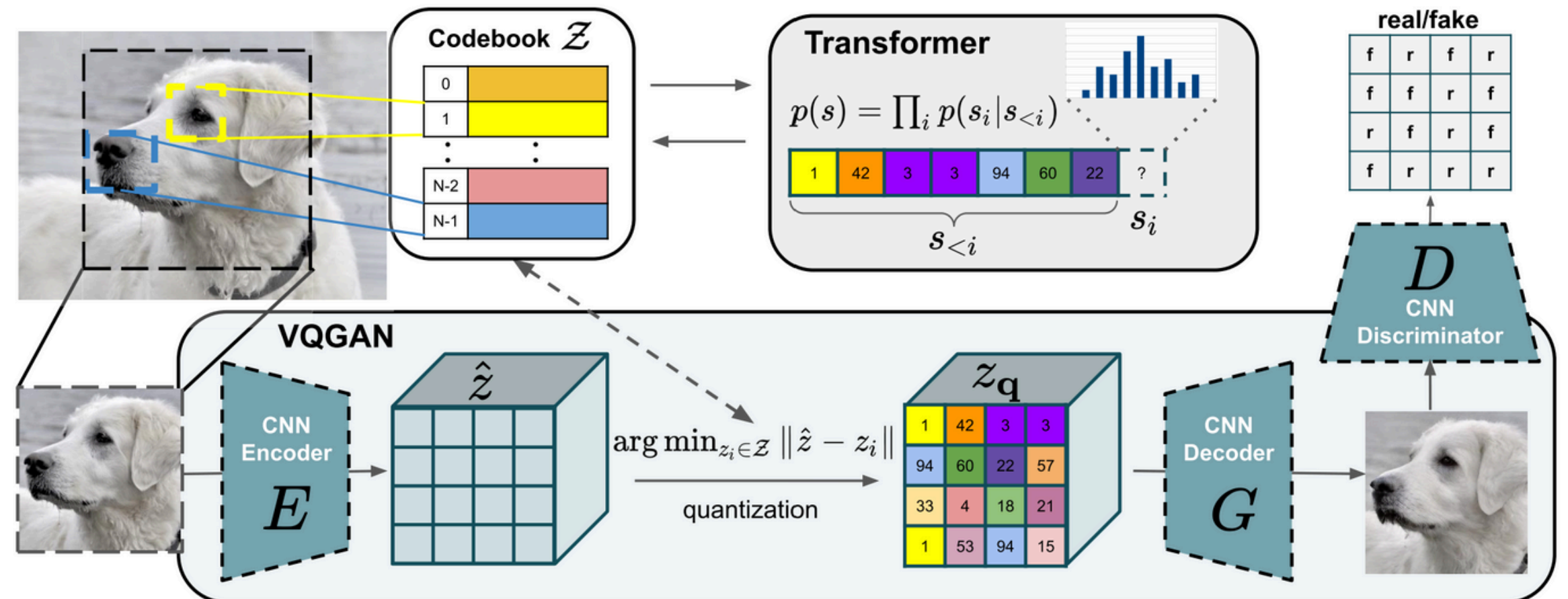
## Only 1.3% trainable params

With use of LoRA modules the total trainable parameters **come down to 1.3%** maintaining inference speed.

# TECHNICAL APPROACH

## WHY THIS APPROACH?

- 1.Task-Specific Fine-Tuning: LOWRANK LoRA **adapts a pre-trained VQVAE model** efficiently, without retraining it fully
- 2.Memory and Computational Efficiency: LOWRANK LoRA **reduces memory usage** by adjusting only a small subset of parameters
- 3.Minimal Data Requirement: Unlike GANs, this method works with unpaired reflection and clean images.
- 4.Preserved Image Integrity: It **removes reflections without degrading** the original image quality, unlike GAN-based methods that can distort the result.

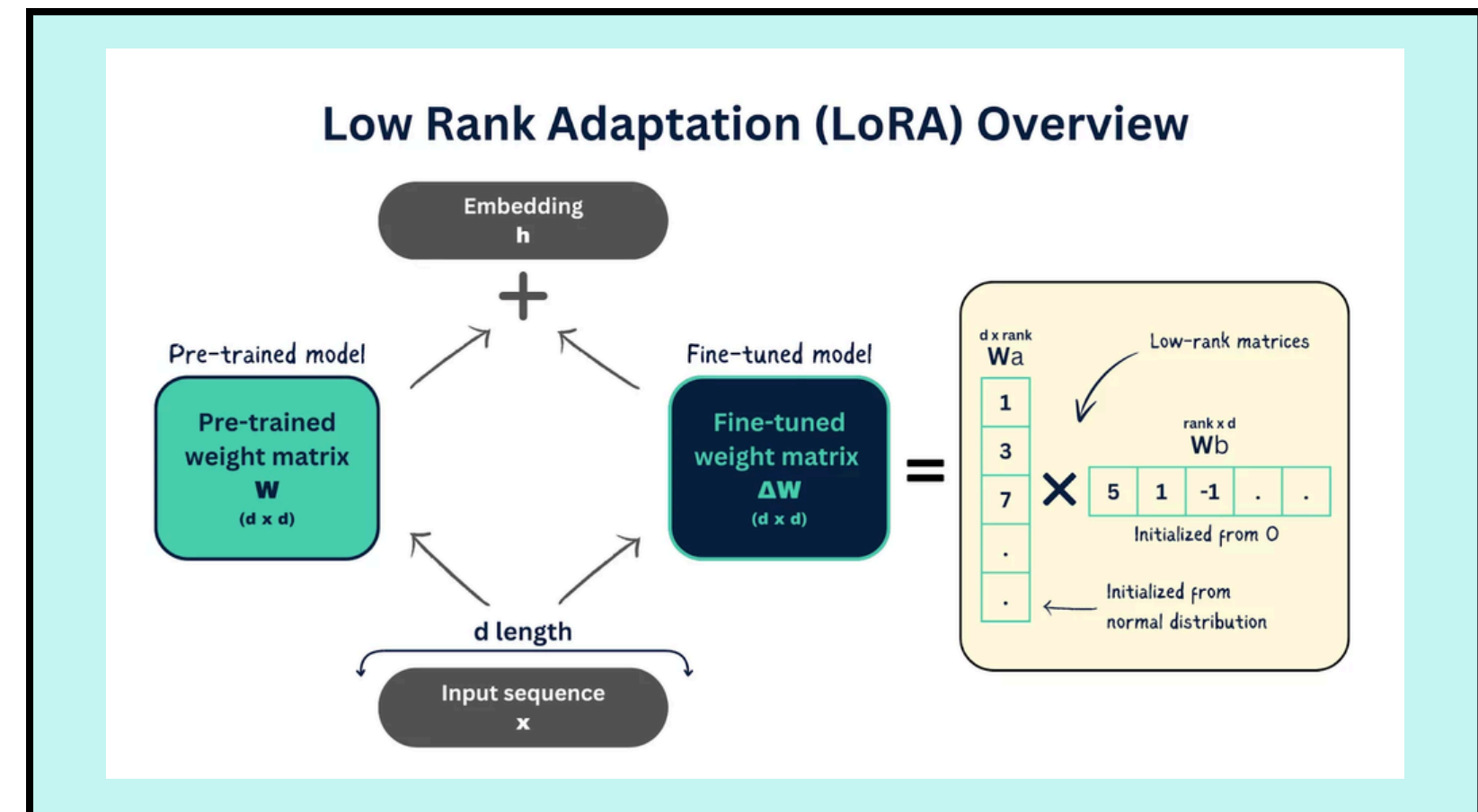


# LoRA

**We empirically observe the intrinsic rank of removing reflection to be less than that of reflection thereby choosing the methodology**

LoRA (Low-Rank Adaptation) is a **technique that allows for task-specific fine-tuning of pre-trained models** without the need to retrain the entire network. This is achieved by adding low-rank matrices to the model that can be adjusted to capture the task-specific features, such as reflection patterns in the image.

In the context of reflection removal, LoRA **helps the VQGAN encoder focus on removing reflections by fine-tuning the latent space representation**. The low-rank matrices are trained to isolate reflection-related features from the image's background and primary content.



# COMPARISON

Feature	Traditional Methods	Deep Learning (CNNs, GANs)	VQVAE + LOWRANK LORA
Training Data	Requires multiple views or precise scene knowledge	Requires large paired datasets	Minimal paired data required
Generalization	Struggles with complex or dynamic scenes	Prone to overfitting or poor generalization	Better generalization to new reflection types
Computational Cost	High computational cost for multiview, complex models	Computationally expensive, especially GANs	Low computational cost, efficient with <b>LOWRANK LORA</b>
Memory Usage	Memory-intensive, especially for stereo or physical models	High memory usage, especially for GANs	Memory efficient, especially due to <b>LOWRANK LORA</b>
Real-Time Performance	Not suitable for real-time applications	Not suitable for real-time unless heavily optimized	Well-suited for real-time applications
Preservation of Image Integrity	Often leads to loss of image quality	May distort or generate unrealistic images	Maintains high image integrity, no distortion