Real-Time Artifact Removal for Passthrough AR

This project explores innovative approaches to enhance the visual fidelity of passthrough augmented reality (AR) systems. By addressing common issues like sensor noise, lighting artifacts, and temporal flickering, we aim to create a more immersive and comfortable AR experience for users. Our focus is on real-time solutions with minimal latency, ensuring practical applicability in dynamic AR environments.

Team VisionAR

- Azizbek — Project Lead
- Azim — Model & Evaluation
- Ikromjon — Data & Integration

The Imperative of Clarity: Why Denoise AR?

Augmented Reality passthrough video, while revolutionary, often struggles with visual imperfections. These issues stem from various sources, including inherent sensor noise, inconsistencies in lighting conditions, and distracting temporal flickering. These artifacts collectively degrade the user experience by compromising visual quality, leading to discomfort, and ultimately breaking the crucial sense of immersion that AR strives to achieve.

Reduced Visual Quality

Distorted images undermine the AR experience.

Decreased User Comfort

Artifacts can cause eye strain and discomfort.

Broken AR Immersion

Unrealistic visuals shatter the illusion of augmented reality.

Our primary goal is to significantly improve AR video quality in real time, ensuring that these enhancements are delivered with exceptionally low latency to maintain a seamless and engaging user interaction.

The Challenge: Real-Time Artifact Removal

Our project directly confronts the complex problem of **real-time artifact removal** specifically tailored for AR passthrough video. This endeavor is fraught with technical challenges, as it demands both high-fidelity image processing and instantaneous delivery to the user.

1

Low Latency

Critical for AR; processing must not introduce noticeable delays (<30 ms).

2

Real-Time Processing

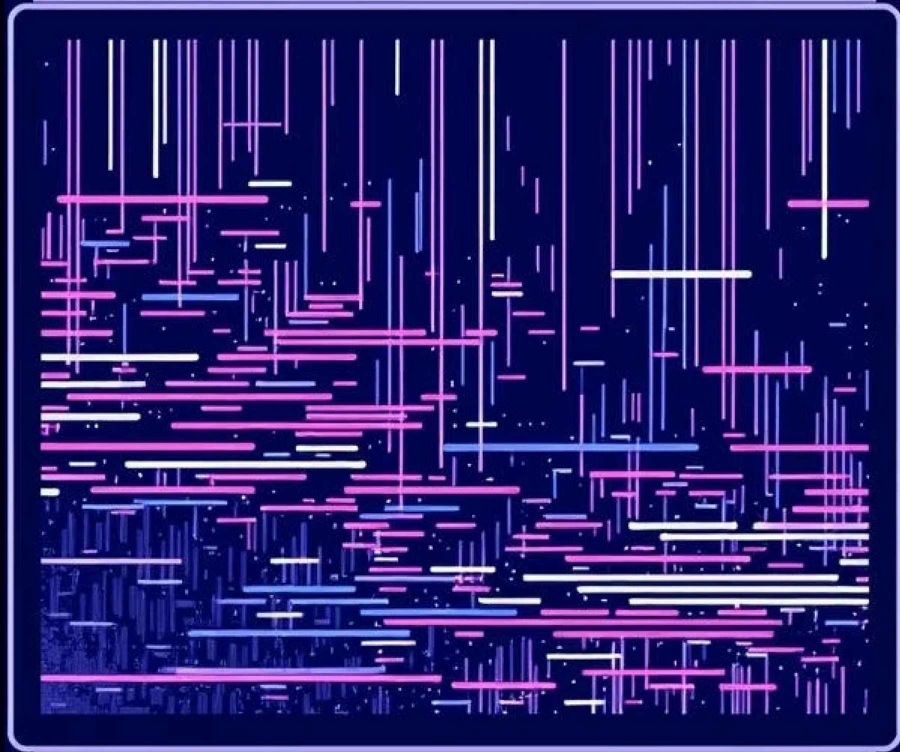
Solutions must operate at video frame rates for live passthrough.

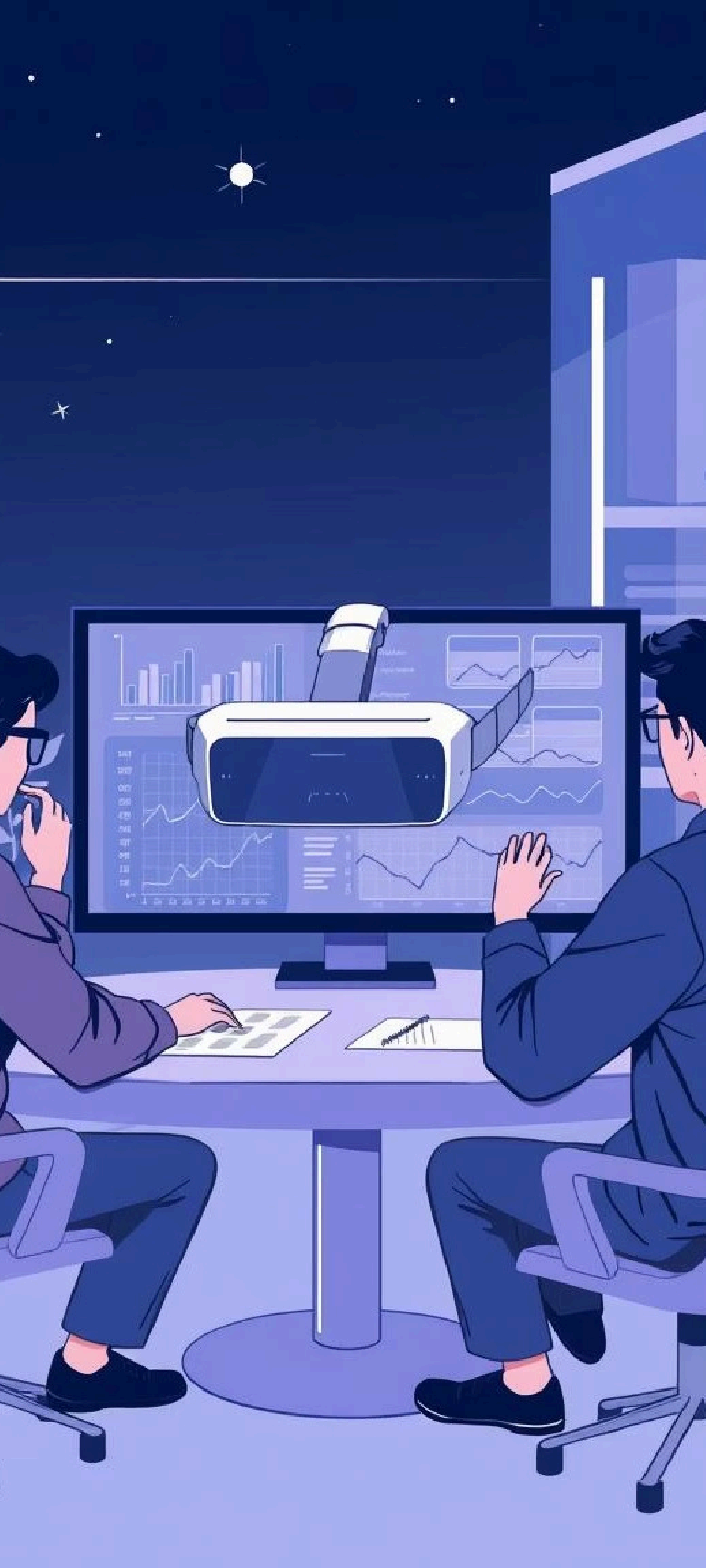
3

Preserve Visual Details

Denoising must not sacrifice essential textures or edges.

Balancing these constraints requires a sophisticated approach that can intelligently distinguish noise from meaningful visual information, all while operating within the strict time budgets of AR systems.





Project Roadmap: Building a Clearer AR Future

To tackle the intricate problem of real-time artifact removal, our project is structured around a series of key objectives designed to progressively build and refine our solution. These objectives guide our development and evaluation process.

01

Baseline Denoising Pipeline

Establish a foundational image processing pipeline.

02

CNN-Based Denoising Model

Develop and integrate an advanced neural network for denoising.

03

Quality and Latency Comparison

Rigorously evaluate the performance of both approaches.

04

Real-Time Feasibility Demonstration

Prove the practical viability of our enhanced AR solution.

Each step is crucial in advancing towards our ultimate goal: a robust, high-quality, and low-latency artifact removal system for passthrough AR.

The Foundation: Our Dataset

A critical component of developing and evaluating our denoising solutions is the selection of an appropriate dataset. We opted for a **single 720p public-domain indoor video from Pexels** due to its ideal characteristics for simulating real-world AR conditions.

- **Natural Noise:** Contains inherent sensor noise common in video feeds.
- **Lighting Variation:** Exhibits diverse lighting scenarios, challenging for consistent image quality.
- **Motion:** Includes dynamic elements and camera movement, crucial for temporal consistency tests.

Why is this dataset suitable? It provides a realistic representation of the visual challenges encountered in actual AR environments, allowing us to accurately assess the effectiveness of our denoising algorithms under relevant conditions. Its lightweight nature also facilitates efficient testing and iteration.

Comparative Analysis: Method Overview

Our approach involves a rigorous comparative analysis of different denoising techniques. We process the input video through distinct pipelines to highlight the impact of each method on visual quality and performance.

Original Video

The raw, unfiltered video input with all its inherent artifacts.

Baseline Denoise

Processed using traditional computer vision filters.

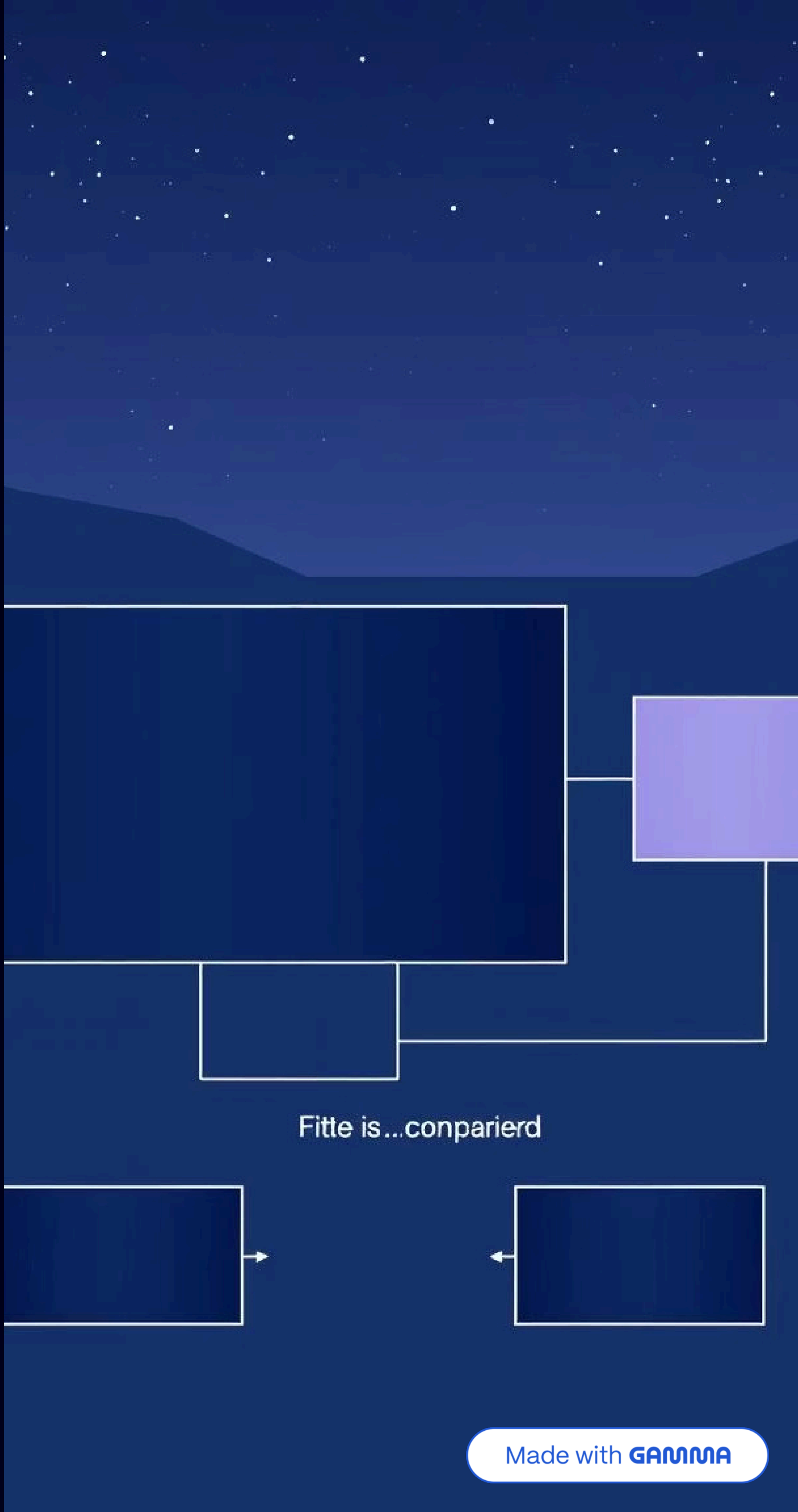
CNN-Based Denoise

Enhanced via a convolutional neural network.

Evaluation & Comparison

Quantitative and qualitative assessment of outputs.

This systematic comparison allows us to clearly demonstrate the advantages and trade-offs of each denoising strategy in the context of AR passthrough.



The Classical Approach: Baseline Denoising

Our baseline denoising method relies on established classical computer vision techniques. This approach serves as a fundamental benchmark against which our more advanced CNN-based method is evaluated. The core components are Gaussian Blur and the Bilateral Filter.

Gaussian Blur

Primarily used for smoothing out general noise.

Bilateral Filter

Offers edge-preserving denoising, attempting to keep details while reducing noise.

Pros:

- **Fast:** Computationally efficient for real-time applications.
- **Simple:** Easy to implement and understand.

Cons:

- **Removes Fine Details:** Can inadvertently smooth out important textures.
- **Appears Blurry:** Often results in an overly soft or blurry image.

The Intelligent Approach: CNN Denoising

The Convolutional Neural Network (CNN) based denoising method represents a significant advancement over traditional techniques. Unlike classical filters, a CNN can learn complex patterns of noise and signal from data, enabling a more intelligent and nuanced approach to artifact removal.

Learns Noise vs. Structure

Discriminates between random noise and meaningful image features more effectively.

Preserves Important Edges

Minimizes the loss of crucial visual details and textures during denoising.

Better Temporal Consistency

Reduces flickering and maintains visual coherence across consecutive frames.

The primary advantage of the CNN approach lies in its ability to yield significantly higher visual quality with less over-smoothing compared to baseline methods, making it ideal for the demanding requirements of AR passthrough.

Quantifying Performance: Evaluation Metrics

To objectively assess the efficacy of our denoising methods, we employ a suite of standard evaluation metrics. These metrics provide a balanced view of both the real-time performance and the visual quality improvements achieved.

FPS

Frames Per Second

Measures real-time processing speed; higher is better.

ms

Latency

Indicates processing delay; lower is better for AR interactivity.

PSNR

Peak Signal-to-Noise Ratio

Quantifies image quality; higher indicates better fidelity.

SSIM

Structural Similarity Index

Assesses perceived image quality based on human visual perception; closer to 1 is better.

By considering these diverse metrics, we ensure a comprehensive evaluation that balances the critical factors of speed and visual quality, which are paramount for a successful AR application.

Promising Outcomes: Initial Results

Our preliminary results demonstrate a clear advantage of the CNN-based denoising approach over the traditional baseline method, especially considering the demanding requirements of real-time AR passthrough.

Metric	Baseline	CNN Method
FPS	31.8	39.25
Latency (ms)	N/A	25.45
PSNR	25.12	28.91
SSIM	0.75	0.84

Key Observation: The CNN not only achieves a higher frame rate and lower latency but also significantly outperforms the baseline in both PSNR and SSIM, indicating superior visual quality while maintaining real-time performance. These results highlight the potential of deep learning for transforming AR visual experiences.