Enabling the Under-Display Camera: Solving Video Quality Using AI Within the ISP

Adriano Scialom, David Huberman and Yoav Taieb
Visionary.ai, Jerusalem, Israel

Abstract
Under-display cameras (UDCs) introduce significant image quality challenges due to optical distortions, increased noise, and reduced light transmission. This paper presents an AI-driven image restoration method designed for real-time execution within an Image Signal Processor (ISP). The approach leverages a recurrent neural network (RNN)-based denoiser with a lightweight architecture optimized for embedded platforms, enabling efficient noise reduction and distortion correction.

We evaluate our method using objective image quality metrics such as Signal-to-Noise Ratio (SNR), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) before and after processing. The AI-enhanced pipeline significantly improves image clarity while operating within power and computational constraints. Comparative analysis with traditional denoising and deblurring algorithms demonstrates this method's advantages in preserving fine details while suppressing structured noise.

Author Keywords
Image Signal Processor; AI algorithms; computer vision; Under-Display Cameras; video quality; image denoising; video denoising.

1. Introduction
The integration of under-display cameras in consumer electronics enables bezel-less designs but significantly degrades image quality due to reduced light transmission, diffraction effects, and spatially varying optical distortions. Conventional ISP-based denoising and deblurring techniques struggle to recover fine details lost in the process. AI-driven approaches have demonstrated improvements in static image restoration, yet real-time video enhancement under computational constraints remains a key challenge.

This paper introduces an AI-based restoration pipeline integrated within the Image Signal Processor (ISP) to correct distortions and enhance video quality in real time. The method leverages spatial and temporal information to improve motion consistency while preserving fine details. For the first time, artificial intelligence has been deployed directly within the ISP of a mass-production laptop featuring an under-display camera, marking a significant milestone in AI-driven image processing. This advancement not only enhances video quality but also paves the way for broader AI adoption across ISP functions, facilitating the widespread use of bezel-less under-display cameras in a variety of display applications.

The implementation of under-display cameras presents unique challenges, primarily stemming from the presence of a physical distortion layer over the sensor. This configuration introduces two major issues: optical distortion and a disrupted noise model. Furthermore, the reduced light transmission through the display exacerbates these problems. Classical algorithms have proven to be insufficient in addressing these complexities, necessitating a novel approach that integrates temporal information to optimize video noise reduction.

2. System Architecture & Methodology
The solution employs a recurrent neural network (RNN)-based video denoising model, optimized for low-latency execution. The core pipeline includes:

2.1 AI-Based Noise and Distortion Correction
- Raw-Domain Processing: Operating directly on raw sensor data before traditional ISP processing.
- Spatio-Temporal Enhancement: Combining information across multiple frames to improve reconstruction.
- Low-Power Execution: Optimized inference on GPU/NPU within power constraints.

2.2 Model Training & Dataset
To train the model, a dataset comprising both real and synthetic UDC images was used. The synthetic data generation process follows an optical simulation-based approach, ensuring realistic degradation patterns. The training dataset accounts for:

- Noise models specific to UDC sensors
- Distortion variations due to different panel architectures
- Temporal consistency constraints for video restoration

3. AI Denoising Network
We propose the use of an AI network deployed in the raw domain, positioned close to the sensor, and specifically trained to accommodate the characteristics of this optical system. This approach effectively reconstructs coherent raw images and videos by combining spatial and temporal data and compensating for missing information. Utilizing a fraction of the neural network inferencing resources of Intel's Lunar Lake processor, the system efficiently performs noise reduction and distortion correction, enabling subsequent processing through the Lunar Lake conventional ISP pipeline.

Our solution leverages recurrent neural networks (RNNs) and additional AI methodologies to capture and optimize temporal and spatial data. Key advantages of this approach include enhanced adaptability to specific optical systems for distortion and noise removal, minimal on-chip Neural Network Processor (NPU), CPU and DDR resource requirements, and seamless integration with existing camera workflows, ensuring system transparency and uninterrupted operation.

ISSN 0097-996X/25/5602-1222-$1.00 © 2025

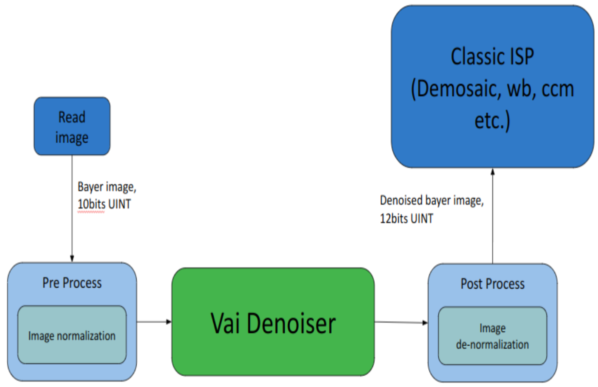

Figure 1. Image pipeline block diagram

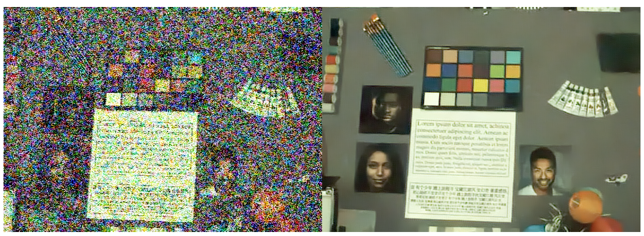

Figure 2. Side by side image examples of original image versus denoised image using raw AI-denoising

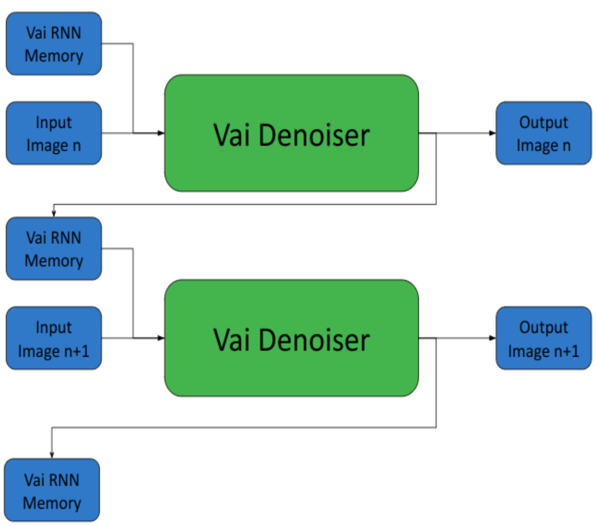

Figure 3. RNN-Based AI Denoising Network Architecture

4. Measured Results

To evaluate our method, we compare performance against conventional denoising and deblurring techniques using SNR, PSNR, and SSIM scores with and without the AI Denoising enabled.

4.1 Signal-to-Noise Ratio (SNR) Improvement

Table 1 presents SNR measurements before and after processing**, demonstrating significant noise reduction without loss of detail.

SNR	Without Denoiser	Traditional Denoiser	With Denoiser
10lux	-7.85	3.84	28.83
30lux	4.24	14.05	30.95
70lux	5.62	14.53	31.63
300lux	9.34	18.04	32.79

4.2 PSNR and SSIM Performance

Our method outperforms traditional algorithms in preserving edge details and texture, as demonstrated in Figure 2 [Placeholder], which compares original vs. AI-enhanced images. Distortion variations due to different panel architectures which are accounted for in the training of the network.

SSIM	Without Denoiser	Traditional Denoiser	With Denoiser
10lux	0.10	0.14	0.34
30lux	0.26	0.32	0.43
70lux	0.28	0.34	0.43
300lux	0.33	0.39	0.45

PSNR	Without Denoiser	Traditional Denoiser	With Denoiser
10lux	14.45	17.74	20.14
30lux	17.26	18.60	18.86
70lux	17.53	18.66	18.79
300lux	18.16	18.77	18.65

5. Temporal Consistency in Video Restoration

One major challenge in video restoration is temporal flicker, where per-frame corrections cause inconsistency across adjacent frames. Our RNN-based model mitigates this effect, as shown in Figure 2, illustrating frame-to-frame stability in real-world sequences.

6. Conclusion

We presented an AI-driven restoration pipeline for under-display camera video enhancement, integrated directly into the ISP. The method effectively reduces noise, corrects distortions, and preserves temporal consistency, enabling real-time execution on embedded platforms.

Our experiments demonstrate significant gains in SNR, PSNR, and SSIM, while ensuring power-efficient processing. Future research will focus on dynamic learning techniques and dataset expansion to enhance robustness across diverse display architectures.