VSRELL: Illuminating the Unseen

A New Approach to Video Enhancement in Low-Light

Presented by the Team from Tianjin University

The Problem We Address

The Dual Challenge

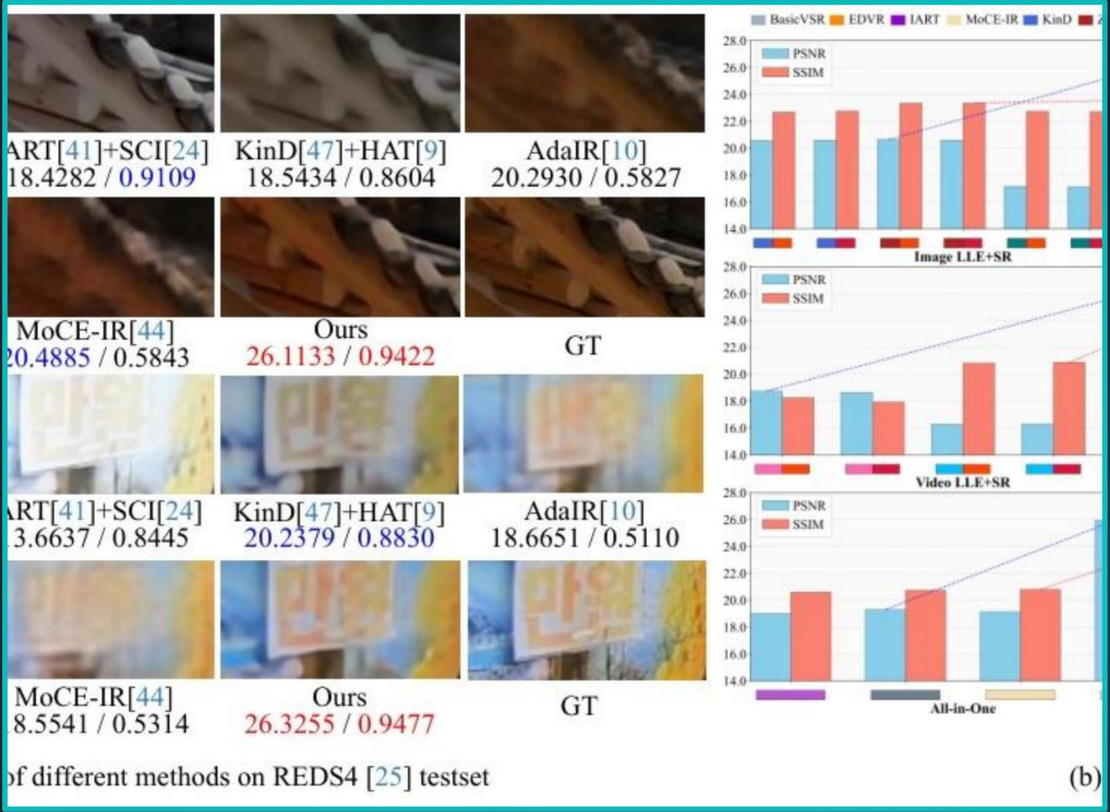
Combining Video Super-Resolution (VSR) and Low-Light Enhancement (LLE) is a complex task due to intertwined degradations.

The Degradation Cocktail

Low-light videos suffer from noise contamination, color distortion, and temporal inconsistency (flickering).

Real-World Impact

Critical for applications like nighttime surveillance, autonomous driving, and mobile video quality.



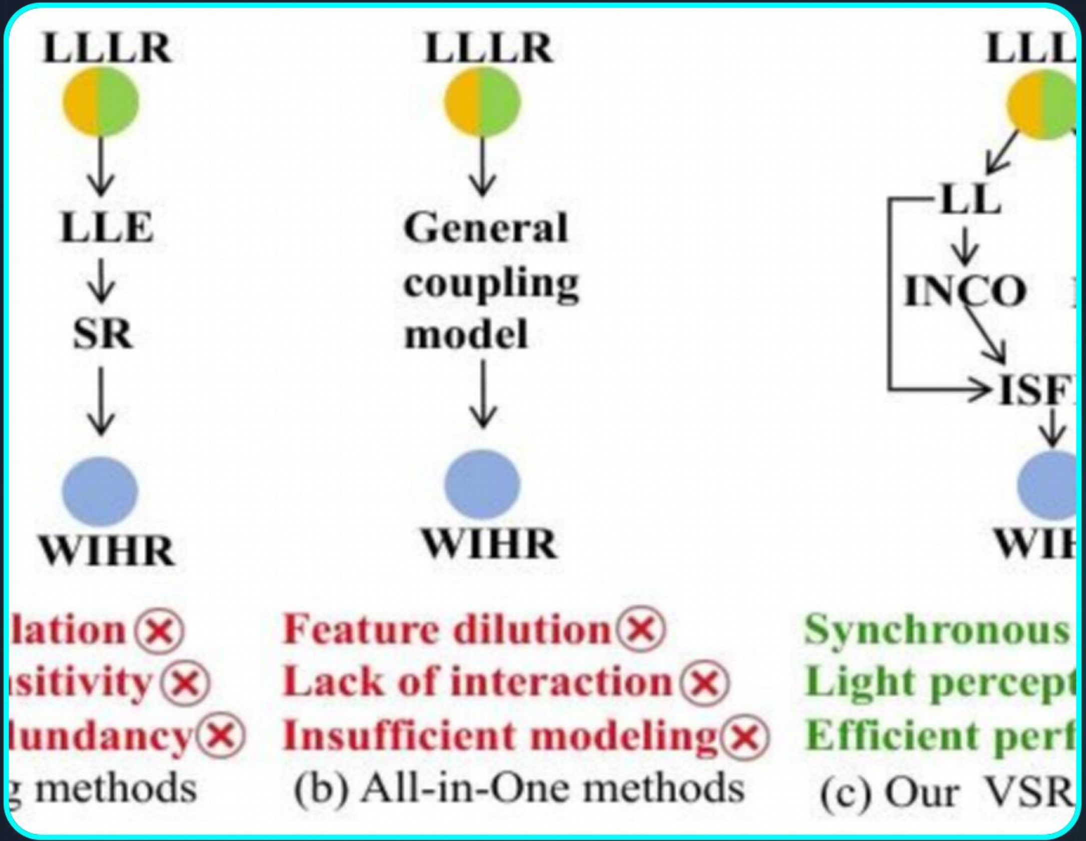
Why Existing Methods Fall Short

Cascaded Methods (LLE → VSR or VSR → LLE)

- **Error Accumulation:** Errors from the first step are amplified in the second.
- **Sequential Sensitivity:** The order of operations matters and can lead to suboptimal results.
- **Parameter Redundancy:** Inefficient use of computational resources.

All-in-One Methods

- **Insufficient Modeling:** Often designed for single-type degradations, not the complex coupling of low-light and low-resolution.
- **Feature Dilution:** May struggle to disentangle different types of degradation.



The Core Scientific Problems

01

Noise-Color Coupling

Noise directly distorts color information in dark regions, making accurate color recovery a significant challenge.

02

Inter-Frame (IF) Information Degradation

Noise obscures fine texture features between consecutive frames, significantly reducing the reliability of motion estimation algorithms.

03

Inaccurate Optical Flow

Low-light environments drastically reduce gradient contrast, leading to frequent motion drift and spatial misalignment.

04

Error Accumulation in Propagation

Initial feature degradation is amplified during long-range temporal propagation, causing severe and distracting temporal flickering artifacts.

Introducing VSRELL: A Unified Approach

A Unified Framework

VSRELL is the first CNN-based method tojointly solveLow-Light Enhancement (LLE) and Video Super-Resolution (VSR) in a single, cohesive end-to-end architecture.

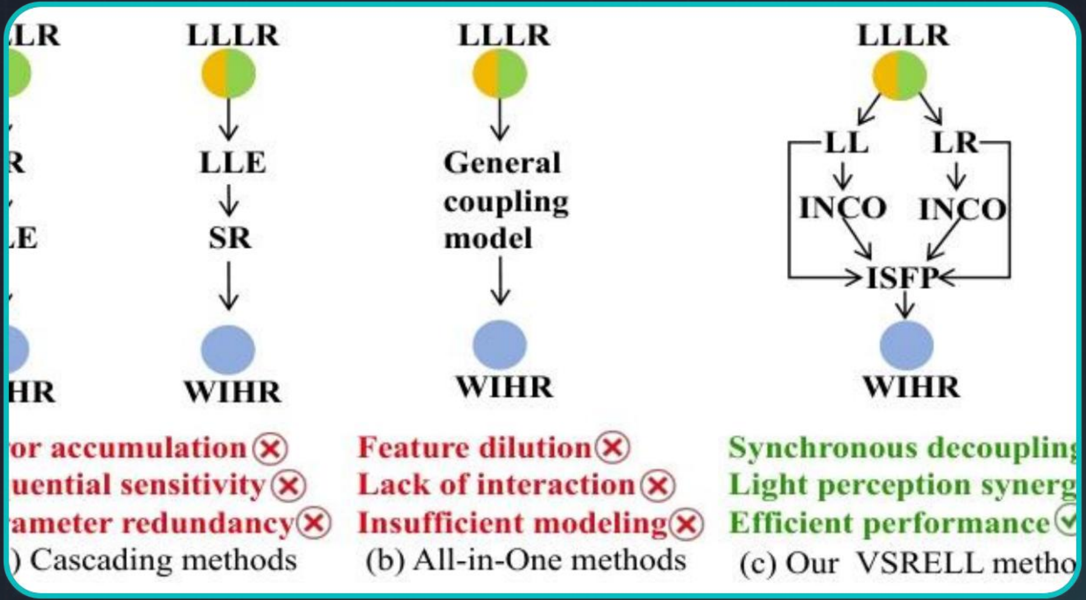
Core Innovations

INCO:Illumination-Noise Co-Optimization — Jointly models and addresses the mutual influence of illumination variations and noise characteristics.

ISFP:Illumination-Sensitive Feature Propagation — Seamlessly integrates illumination information into feature propagation to preserve fine details.

Our Primary Goal

To achievesynchronous decouplingof highly intertwined degradations (noise, blur, low-light) for both computation-efficient and high-quality visual restoration.



Methodology & Performance

The diagram above compares VSRELL against state-of-the-art methods, highlighting our unique advantages in:

"Synchronous Decoupling" | "High Efficiency"

Deep Dive: The INCO Module

Illumination-Noise Co-Optimization

Joint Modeling

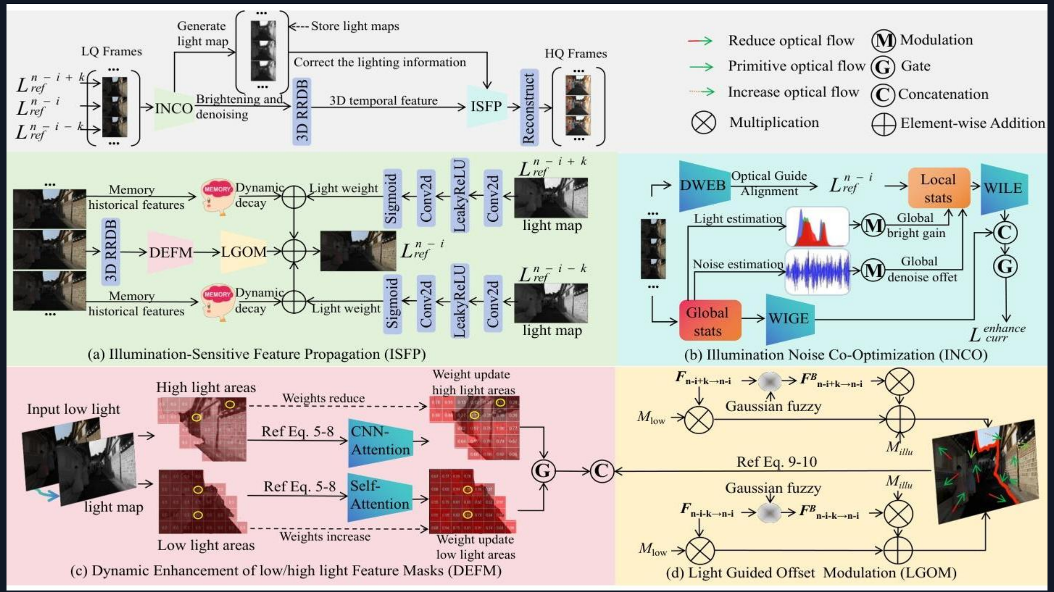
Treats illumination enhancement and noise suppression as a single, coupled problem for optimal results.

Dynamic Window Partitioning

Analyzes sequences of frames to accurately model global illumination trends and noise statistics.

Global-Local Feature Fusion

Combines overall scene context with fine-grained local details for a balanced enhancement effect.



Adaptive Brightness Adjustment

Dynamically adjusts pixel gains to avoid common artifacts like over-exposure or under-exposure.

Deep Dive: The ISFP Mechanism

Illumination-Sensitive Feature Propagation

Illumination-Aware Alignment

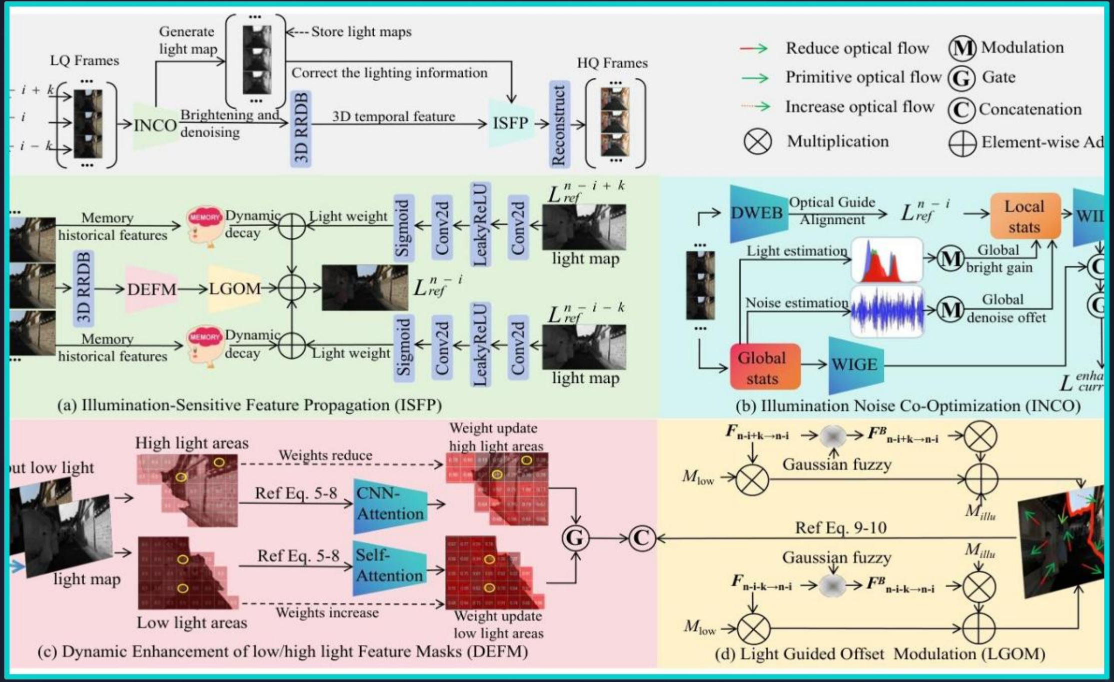
Adjusts motion estimation based on illumination, being more flexible in dark regions to avoid alignment errors.

Optical Flow-Adaptive Smoothing

Reduces noise in dark areas where signals are weak, while preserving fine details in bright, reliable regions.

Dynamic Memory Decay

Intelligently weights features from past frames, downplaying unreliable frames to break the error accumulation chain.



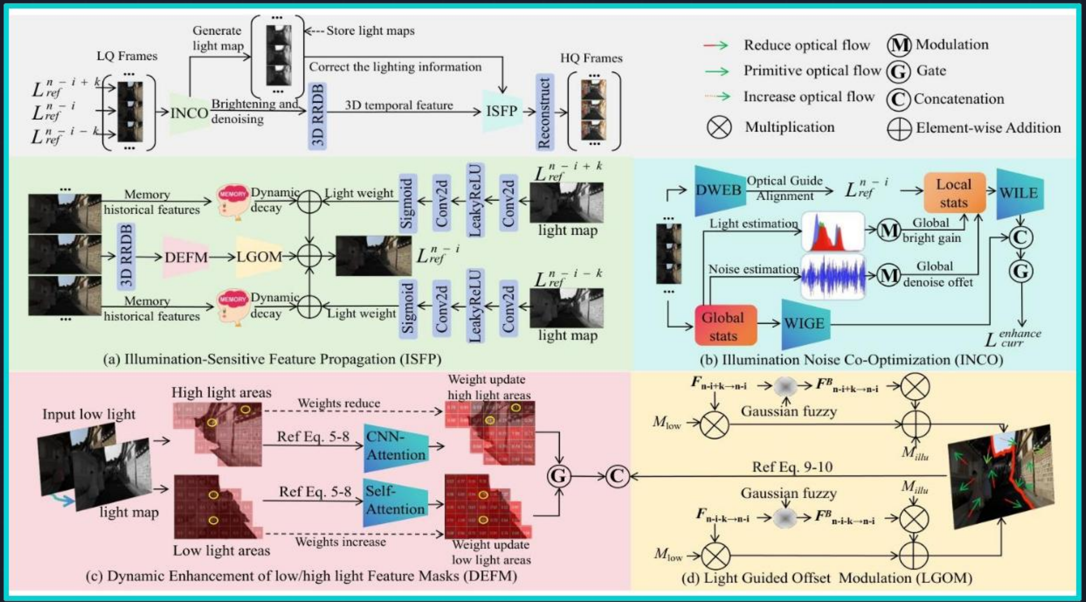
VSRELL Architecture

Highlighting the ISFP (Illumination-Sensitive Feature Propagation) Module

The Complete VSRELL Architecture

End-to-End Pipeline

A single, unified network designed to handle the entire enhancement task in one streamlined pass, eliminating data fragmentation and inefficiency.



Final Output

The result is a high-quality video stream with significantly enhanced brightness, upscaled resolution, and strong temporal consistency across all frames.

Component Synergy: INCO + ISFP

The**INCO**(Input Correction) module prepares raw video features by cleaning up noise and artifacts. It then hands off the refined data to the**ISFP**(Inter-Frame Feature Propagation) module, which ensures these high-quality features are accurately and smoothly propagated across time, maintaining spatial and temporal coherence.

How We Tested VSRELL

DATASETS

- REDS (training & testing)
- Vid4 & UDM10 (generalization testing)

COMPETITORS

- Comprehensive list of SOTA methods
- Cascaded combinations (e.g., SwinIR+KinD)
- All-in-One models (e.g., AdaIR, PromptIR)

METRICS

- Peak Signal-to-Noise Ratio (PSNR)
- Structural Similarity Index (SSIM)

*Higher is better**

Quantitative Results: VSRELL is the New Leader

SOTA Performance

VSRELL outperforms all compared methods on REDS4, Vid4, and UDM10 datasets, setting new benchmarks.

Significant Gains

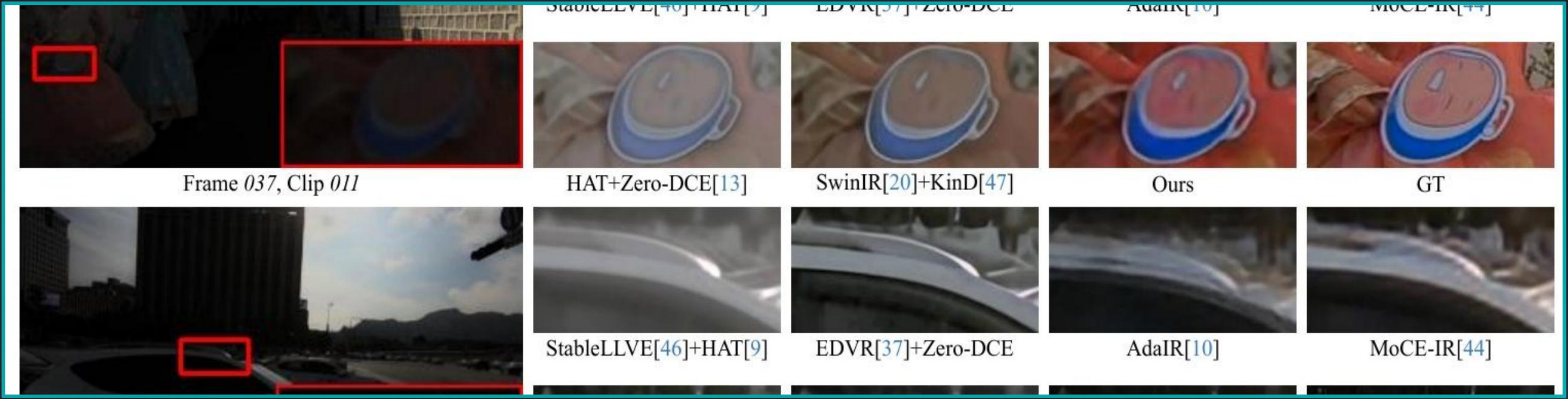
Achieves substantial improvements across the board in both PSNR and SSIM, the gold-standard image quality metrics.

Strong Generalization

Maintains high performance on completely unseen data, demonstrating excellent robustness to different distributions.

Method	REDS4 (PSNR / SSIM)	Vid4 (PSNR / SSIM)	UDM10 (PSNR / SSIM)
Method A (SOTA)	32.51 / 0.8962	31.89 / 0.8815	30.12 / 0.8643
VSRELL (Ours)	33.72 (+1.21) / 0.9105 (+0.0143)	33.15 (+1.26) / 0.8980 (+0.0165)	31.50 (+1.38) / 0.8801 (+0.0158)

Qualitative Results: A Clear Difference



Superior Visual Quality

VSRELL produces sharper details, more natural colors, and fewer artifacts compared to competing methods.

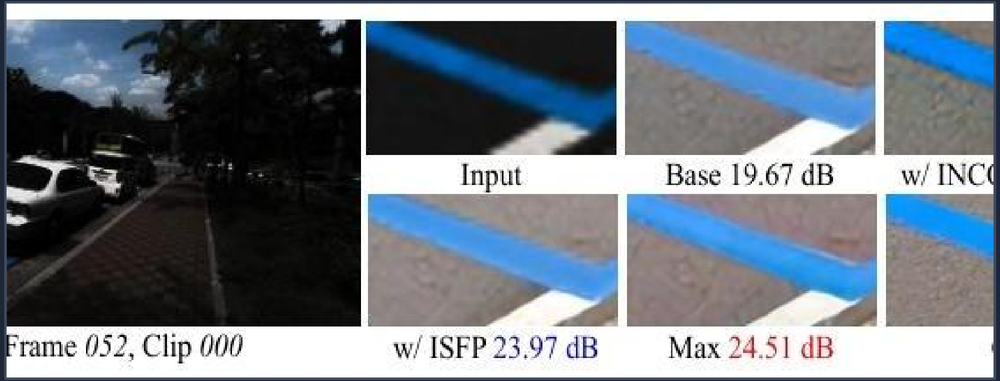
Temporal Stability

Effectively suppresses flickering and maintains structural integrity in dynamic regions across frames.

Realistic Enhancement

Strikes a perfect balance to avoid over-smoothing or creating artificial-looking, "plastic" results.

Ablation Studies: Proving Our Components Work

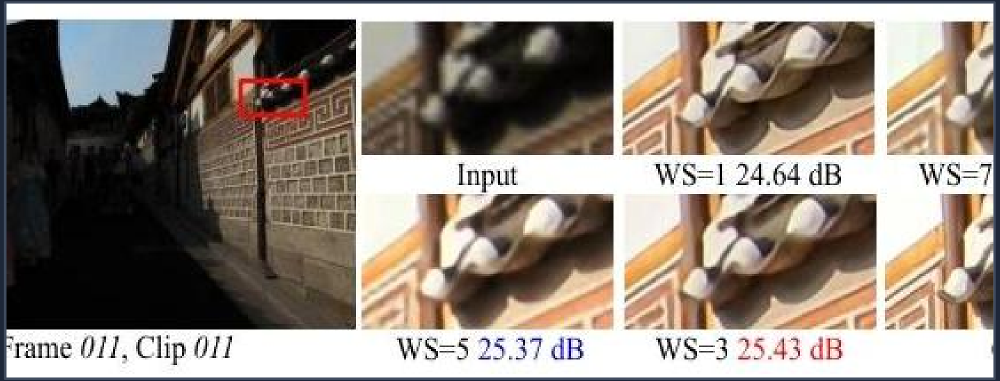


INCO & ISFP are Essential

Ablation tests reveal that removing either INCO or ISFP module results in a significant performance degradation of over 0.8 dB in PSNR, demonstrating their critical role.

Component Synergy

The integration of INCO and ISFP yields the best qualitative and quantitative results, proving the effectiveness of their combined design.



Optimal Window Size

We systematically tested window sizes ranging from 1 to 7 frames. The data consistently showed that a window size of 3 frames (WS=3) achieves the best balance between computational efficiency and model accuracy.

This finding informed our final architecture, ensuring it is both lightweight and powerful.

Conclusion

A New Baseline

VSRELL provides a simple yet effective baseline for the challenging task of joint low-light video enhancement and super-resolution.

Proven Effectiveness

Both quantitative and qualitative results demonstrate VSRELL's superiority over existing methods.

Technical Contribution

The INCO and ISFP modules offer novel solutions to the problems of noise-color coupling and error accumulation in video restoration.

Future Work

Handling More Complex Degradations

Extend the current model to address more extreme and diverse multi-degradation scenarios. We aim to tackle combinations of heavy noise, motion blur, high compression artifacts, and low-light conditions that remain challenging for state-of-the-art restoration methods.

Efficiency & Deployment

Optimize the model architecture for computational efficiency to enable seamless deployment on mobile and edge devices. The goal is to achieve real-time inference speeds without a significant drop in accuracy, unlocking practical use cases for everyday photography.

THANK YOU!

Questions & Answers

FEEL FREE TO ASK ANYTHING