IEEE MMSP 2024
October 2 - 4, 2024, Purdue University, West Lafayette, Indiana, U.S.A.

Diffusion-based Bit-depth Expansion

Riyu Lu[1, 2], Lingyu Zhu[1], Baoliang Chen[1], Xiaopeng Fan[2], Shiqi Wang[1]

[1] City University of Hong Kong
[2] Harbin Institute of Technology

Introduction

Background

- **Bit-depth:** Key factor in visual fidelity, crucial in fields like medical imaging, video streaming, and computer vision.
- **Human Perception:** Requires higher bit-depths to capture the full range of colors and luminance.
- **Modern Display Technologies:** Support 10-bit or 12-bit depth (HDR monitors, high-end TVs), surpassing the traditional 8-bit standard.

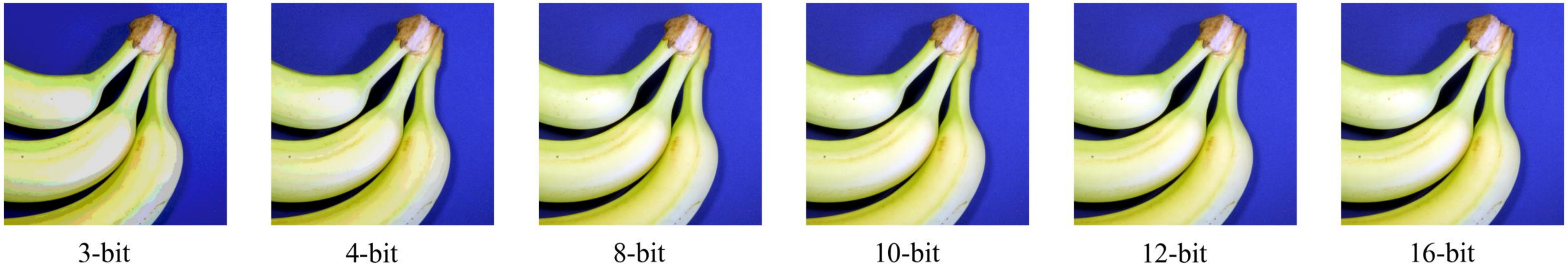

Fig 1. A visual comparison of 8-bit, 10-bit, and 12-bit images to highlight differences in color depth and visual quality.

Introduction

Motivation

- **Issue:** Image processing often quantizes data to 8-bit per channel for compatibility, leading to color loss and banding artifacts.
- **Solution:** Bit-depth expansion techniques reconstruct 8-bit images into high-bit-depth versions, enhancing visual quality and color precision.

Related Works

Traditional Methods

- Rely on handcrafted rules
- Prone to artifacts like false contours and banding
- Poor quality

Deep Learning-based Methods

- Leverage neural networks for color and brightness accuracy
- Improved quality but face challenges:
    - New visual artifacts
    - High computational cost

Related Works

Diffusion Models: A New Frontier

- Diffusion models: Revolutionizing image generation, denoising, and super-resolution
- Exceptional in texture synthesis and image restoration
- Potential for bit-depth expansion remains unexplored
- First study to apply diffusion models for bit-depth expansion

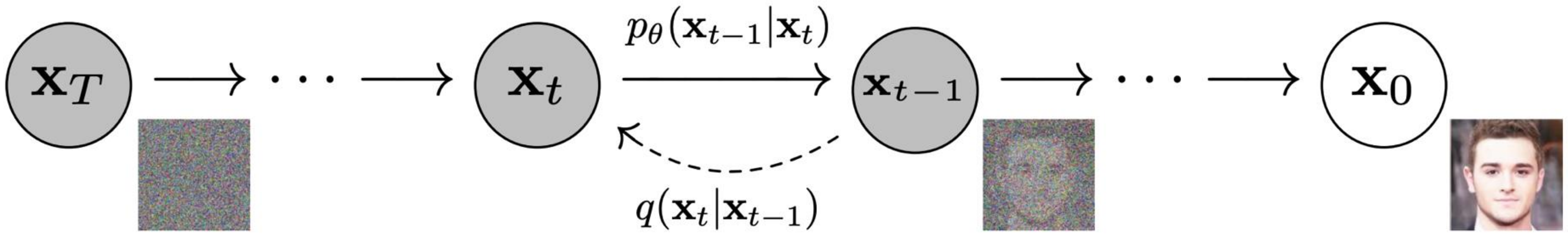

Fig 2. A simplified diagram showing the diffusion model process.[1]

[1] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, *33*, 6840-6851.

Wavelet-based Diffusion Model for Bit-depth Enhancement

Wavelet Decomposition

- 2D Discrete Wavelet Transformation
- Advantage: Targeted processing of components
- Low-frequency: Diffusion model
- High-frequency: Specialized filtering

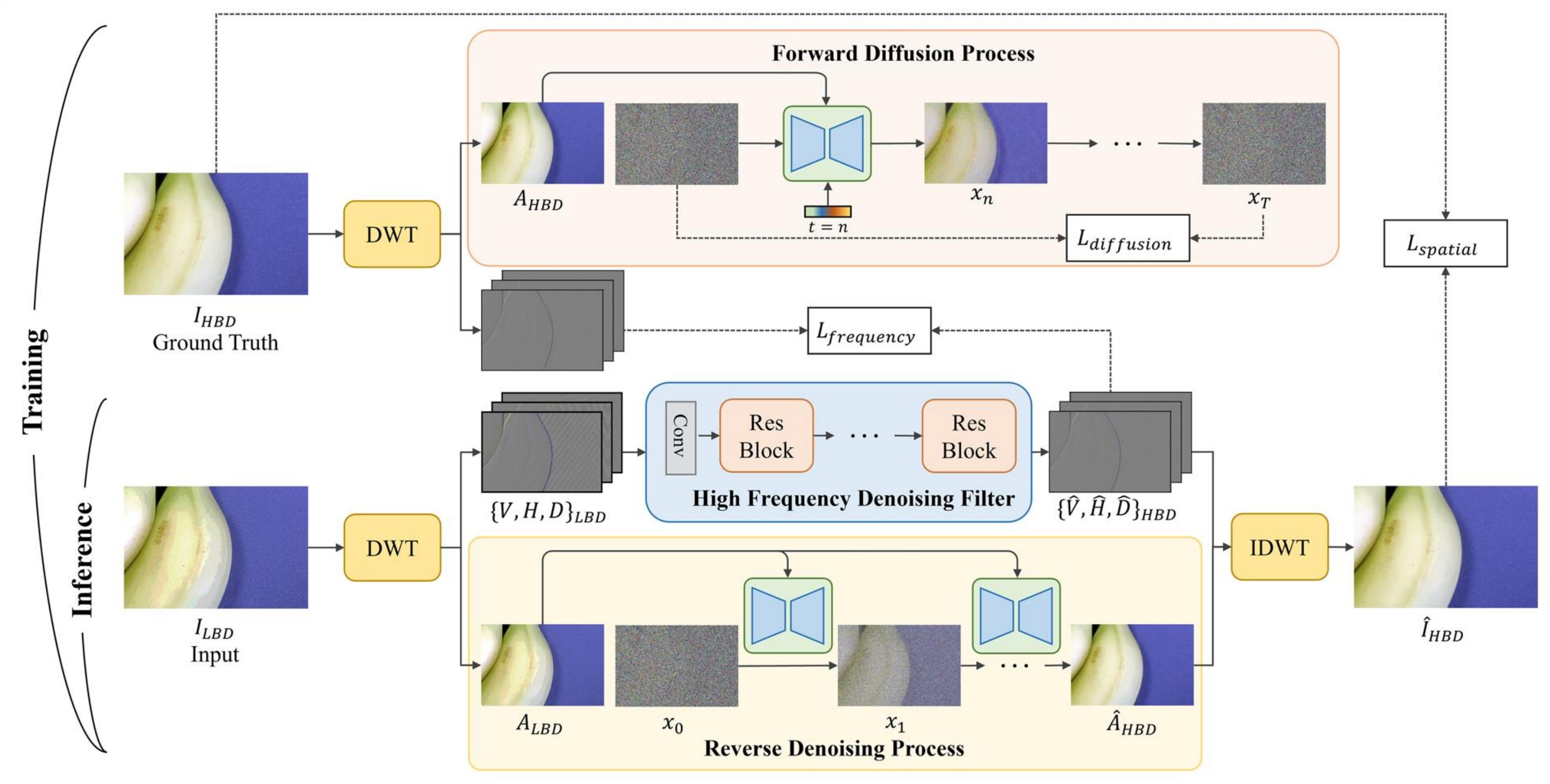

Fig 3. Overall architecture of the proposed bit-depth expansion method.

Wavelet-based Diffusion Model for Bit-depth Enhancement

Low-Frequency Component Processing

- Forward Diffusion: Adds Gaussian noise
- Reverse Denoising: Uses low-bit-depth image as conditional input
- Result: Restores structure and shading of the image

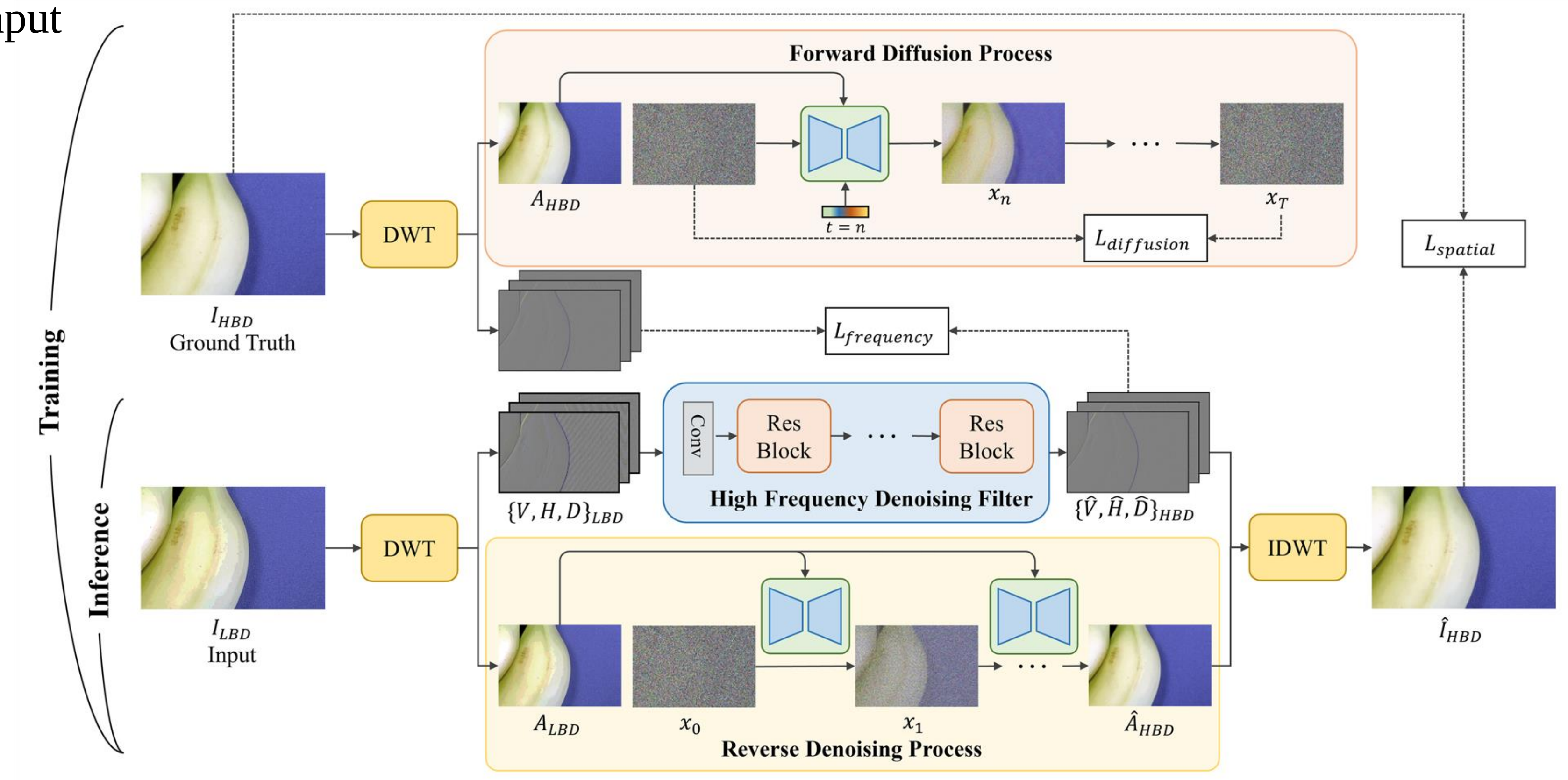

Fig 3. Overall architecture of the proposed bit-depth expansion method.

Wavelet-based Diffusion Model for Bit-depth Enhancement

High-Frequency Component Processing

- Reduces noise & artifacts (e.g., false contours)
- Preserves fine details, textures, sharp edges

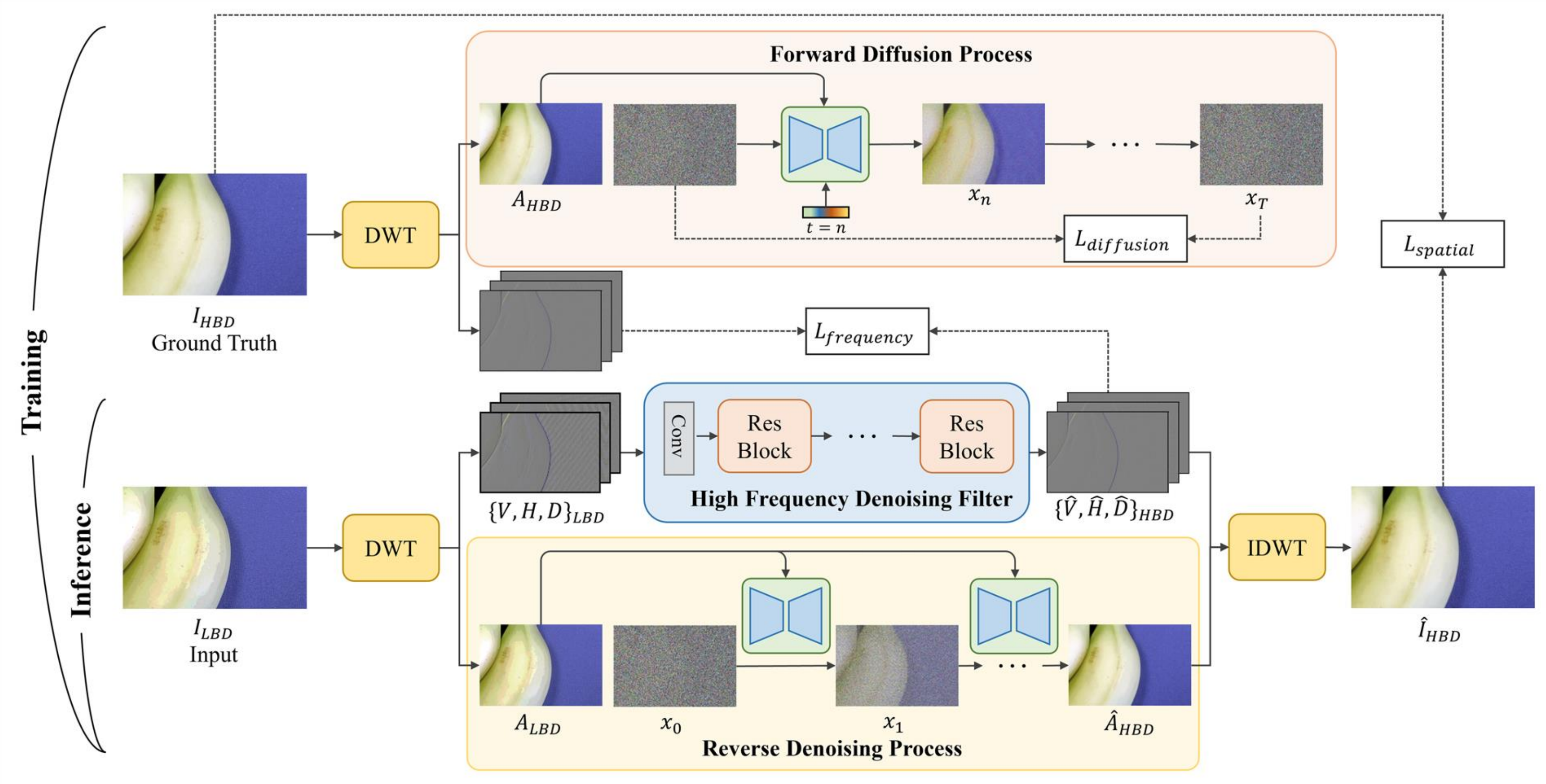

Fig 3. Overall architecture of the proposed bit-depth expansion method.

Final Reconstruction

- Combine low- and high-frequency components
- Inverse Wavelet Transformation
- Output: High-bit-depth image with smooth transitions & preserved details

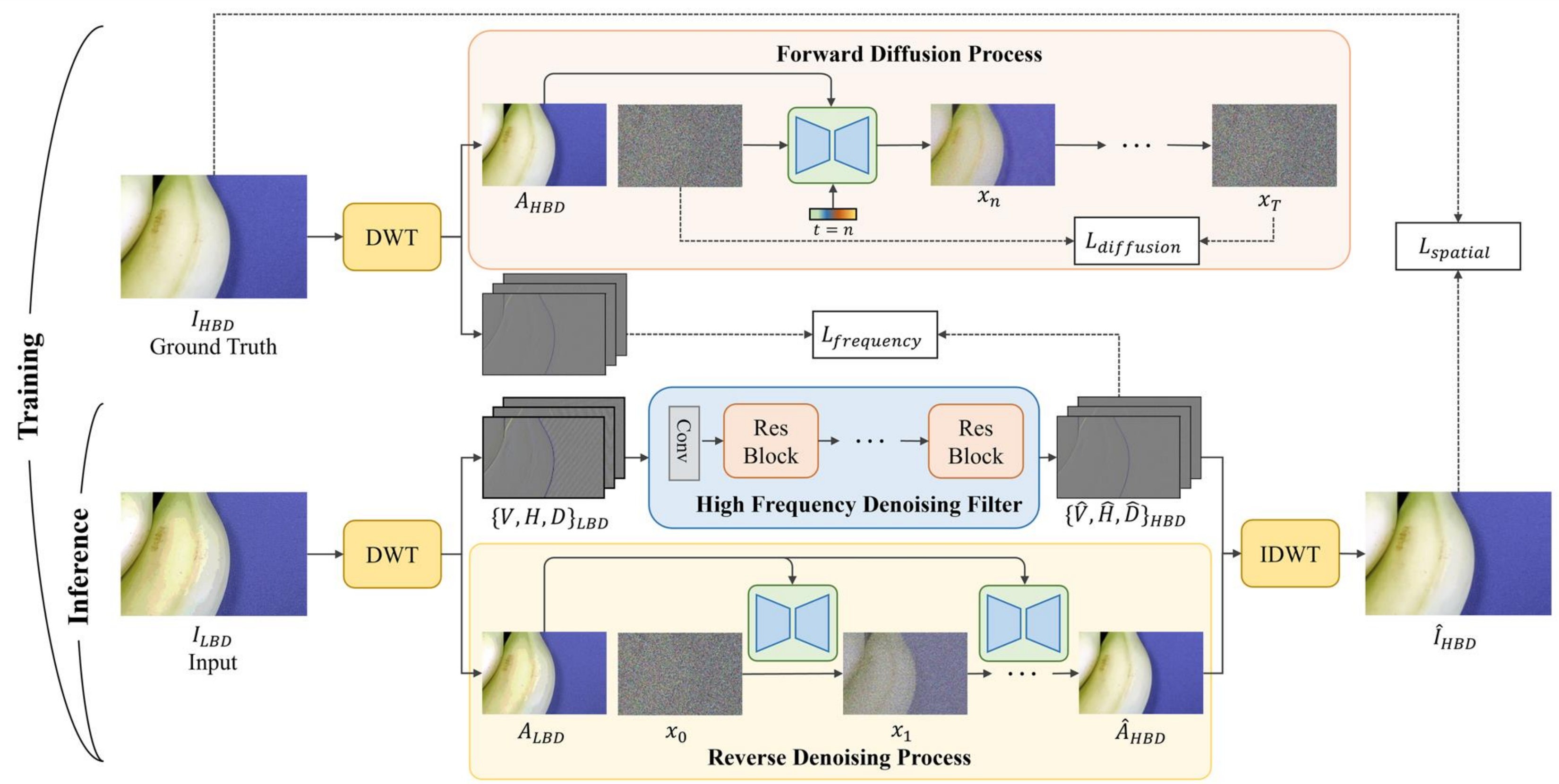

Fig 3. Overall architecture of the proposed bit-depth expansion method.

Experimental Results

Dataset and Training

- 2,000 16-bit images
- Sources: Sintel (animated), MIT-Adobe FiveK (natural scenes)
- Test set: 50 Sintel images

Quantitative Results

Table I. Quantitative results on SINTEL test dataset. The highest scores in each setting are **bolded**, and the second-highest scores are <u>underlined</u>.

Method	Sintel 4→8		Sintel 4→12		Sintel 4→16		Sintel 6→16		Sintel 8→16	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Input	29.16	0.8864	28.79	0.8844	28.77	0.8843	40.81	0.9857	52.86	0.999
MRC	33.95	0.9141	33.79	0.9126	33.78	0.9126	46.85	0.9903	59.31	0.9993
CRR	34.36	0.9389	33.83	0.9352	33.8	0.9348	46.02	0.9864	57.41	0.9981
CA	35.71	0.9444	35.52	0.9438	35.5	0.9436	46.96	0.9896	57.85	0.9988
ACDC	34.6	0.9074	34.64	0.9077	34.64	0.9077	46.66	0.9858	58.7	0.9989
IPAD	35.86	0.9457	35.78	0.9452	35.76	0.9451	47.62	0.9902	58.62	0.9989
BE-CNN	35.68	0.9566	35.71	0.9578	35.71	0.9578	49.74	0.9926	54.78	0.9989
BitNet	39.34	0.9701	39.49	0.9719	39.49	0.9719	49.68	<u>0.9954</u>	57.55	0.9989
BE-CALF	<u>39.91</u>	<u>0.9737</u>	<u>39.98</u>	<u>0.9752</u>	<u>39.98</u>	<u>0.9752</u>	**51.14**	0.994	<u>59.51</u>	<u>0.9993</u>
Ours	**41.05**	**0.9778**	**40.84**	**0.9778**	**41.18**	**0.9787**	<u>50.88</u>	**0.9966**	**59.77**	**0.9995**

Experimental Results

Qualitative Results

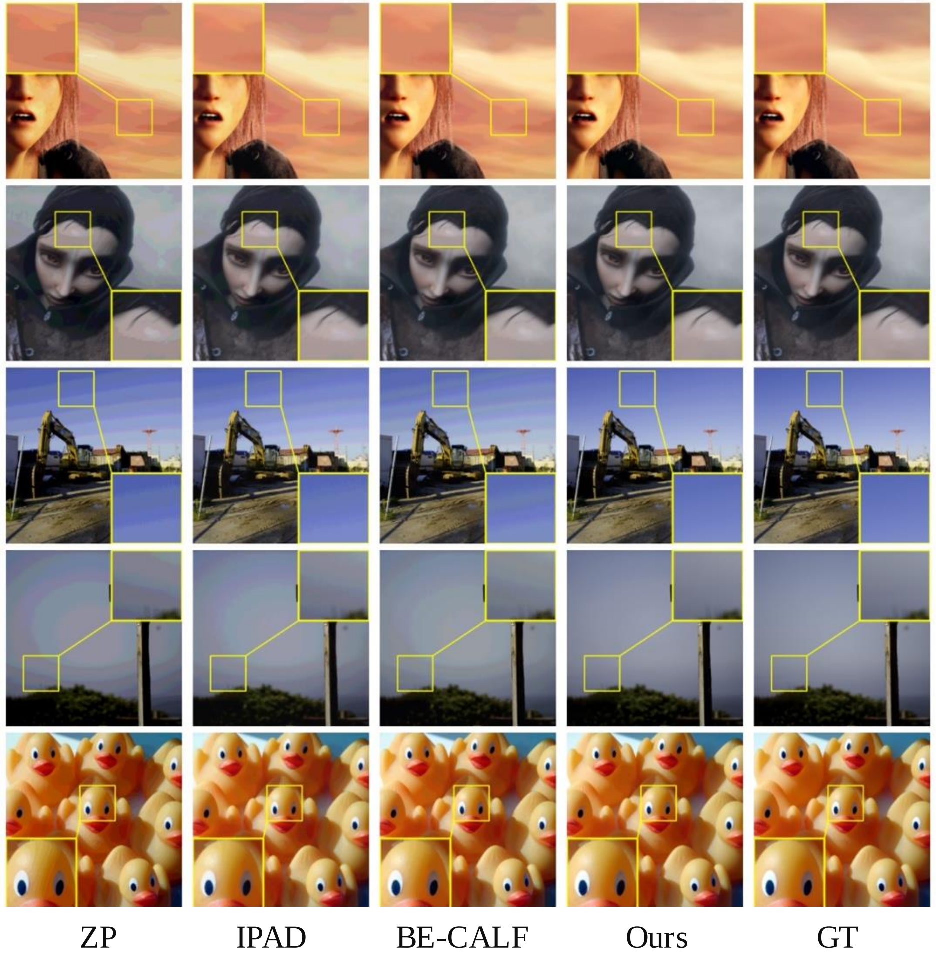

Fig 4. Qualitative comparison on the animated images and natural images for 4-bit to 16-bit recovery.

Computational Analysis

Table II. Complexity Results of Different Methods.

Method	Inference Time (s)	Model Complexity (M)
MRC	594.2	-
CRR	102.74	-
CA	118.3	-
ACDC	1205.91	-
IPAD	8.99	-
BE-CALF	0.34	5.2
Ours	0.24	20.76

Conclusion and Future Insights

Future Applications

Diffusion Models:

- Image restoration: denoising, super-resolution, inpainting
- Compression artifact removal
- Useful for mobile/cloud streaming

Compression Pipelines:

- Efficient storage & transmission of UHD/HDR video
- Integrate with codec for improved coding efficiency
- Balance compression efficiency & quality

Diffusion-based Bit-depth Expansion

Thank you!