Activating Capability of Linear Attention for Image Restoration

Yubin Gu1, Yuan Meng1, Jiayi Ji12, Xiaoshuai Sun1
1Key Laboratory of Multimedia Trusted Perception and Efficient Computing, Ministry of Education of China, Xiamen University, China
2National University of Singapore, Singapore

Introduction

Image Restoration (IR) is essential for recovering clear images from degraded inputs, supporting downstream vision tasks. While CNN-based models are efficient, they lack global context modeling. Transformer-based IR methods address this but suffer from high computational cost due to quadratic attention.

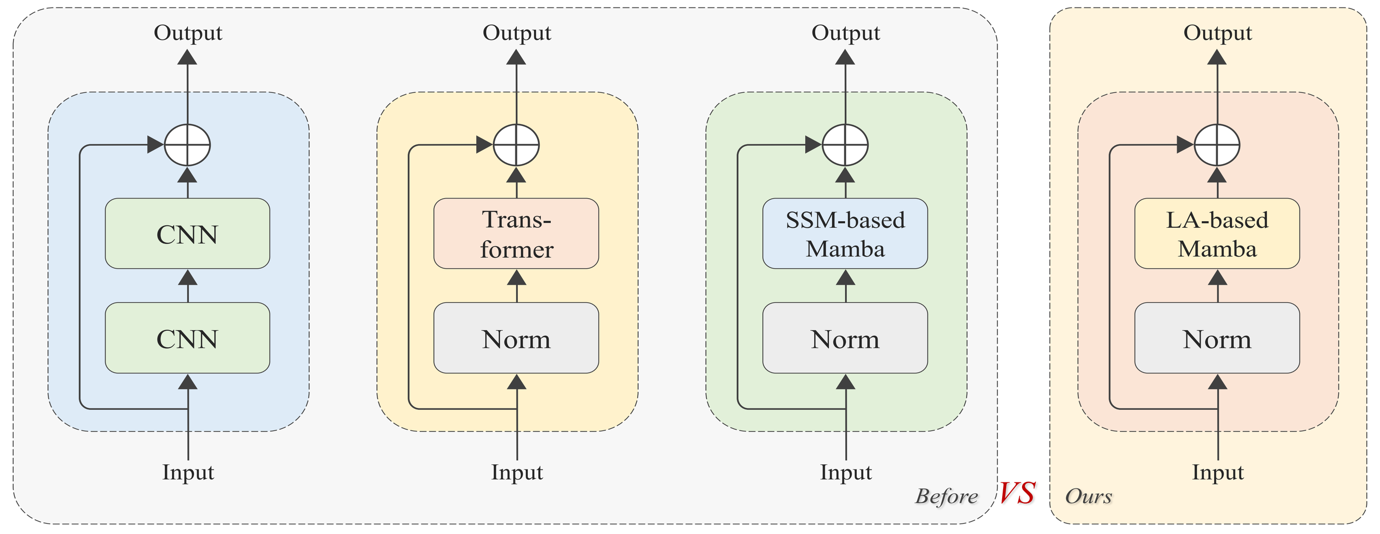
State-space models like Mamba offer efficient alternatives for sequence modeling but face challenges when applied to 2D images, flattening disrupts spatial locality, and unidirectional recurrence ignores multi-directional dependencies.

To overcome these issues, we propose ACL, a novel IR architecture that replaces Mamba's state-space module with Linear Attention, enabling efficient, multi-directional global feature modeling with linear complexity. We design LAMA, a core module built on LA to capture long-range dependencies, and MDC, a lightweight multi-scale dilated convolution module to enhance local detail restoration.

ACL forms an encoder-decoder structure that balances global and local modeling. On deblurring and deraining benchmarks, ACL achieves competitive or superior performance compared to SOTA Transformer models, with significantly fewer parameters and lower inference cost.

Main Contributions

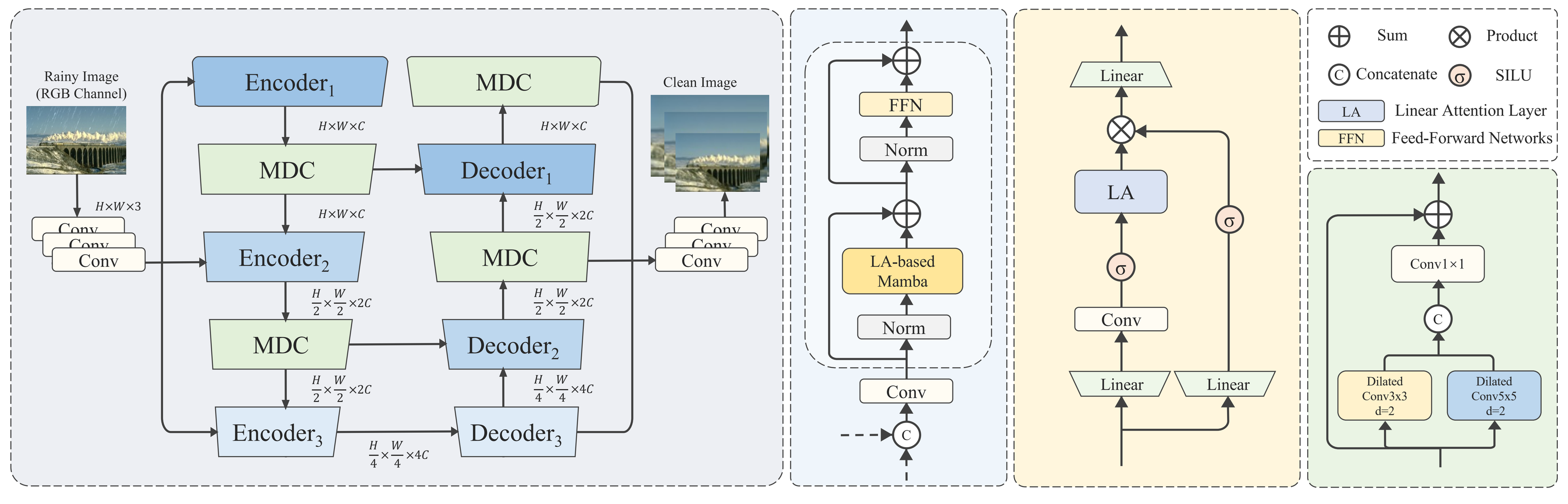
- We propose ACL, combining Linear Attention with the Mamba framework for efficient global modeling in image restoration.
- We introduce LAMA for global dependencies and MDC for local detail enhancement.
- ACL achieves state-of-the-art performance with lower computational complexity.



Weak in Global Views (a) CNN-based Block
Weak in Efficiency (b) Transformer-based Block
Weak in Multi-Views (c) SSM-based Mamba Block
Balance (d) LA-based Mamba Block

Figure 1. Basic modules for different mechanisms

Overall Pipeline



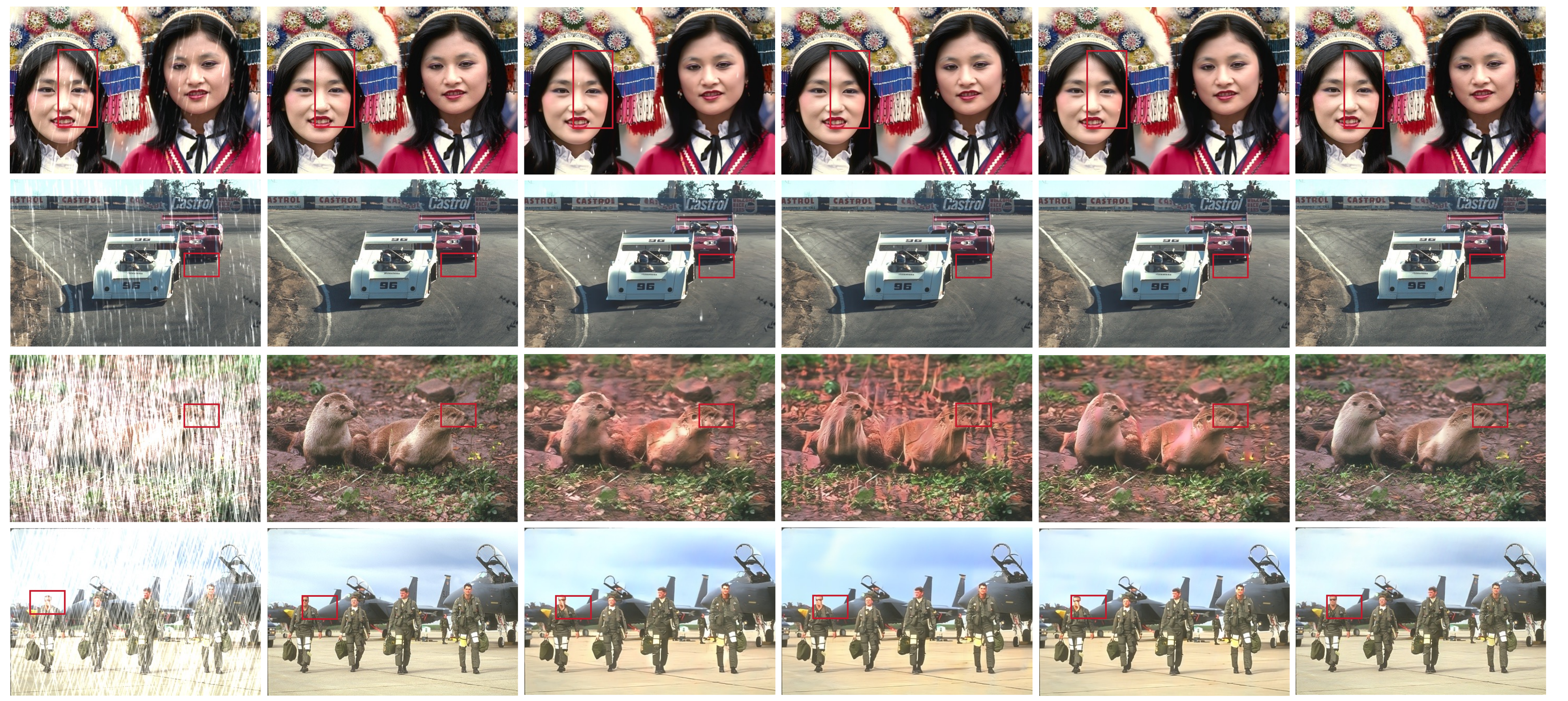
(a) Overall Pipeline (b) Encoder and Decoder (c) LA-based Mamba (d) Multi-Dilated Convolutions Module

Figure 2. The Pipeline of our proposed IR model, including LAMA and MDC modules.

Result

Methods	Restormer		MAXIM		MPRNet		IRNeXt		MambaIR		ACL (Ours)	
Dataset	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rain100L	38.99	0.978	38.06	0.977	37.84	0.959	38.24	0.972	38.78	0.977	**39.18**	**0.983**
Rain100H	31.46	0.904	30.81	0.903	30.41	0.874	31.64	0.902	30.62	0.893	**32.22**	**0.920**
Average	35.23	0.941	34.44	0.940	34.13	0.917	34.94	0.937	34.70	0.935	**35.70**	**0.952**

Table 1. Quantitative comparison results on the Rain100L and Rain100H.

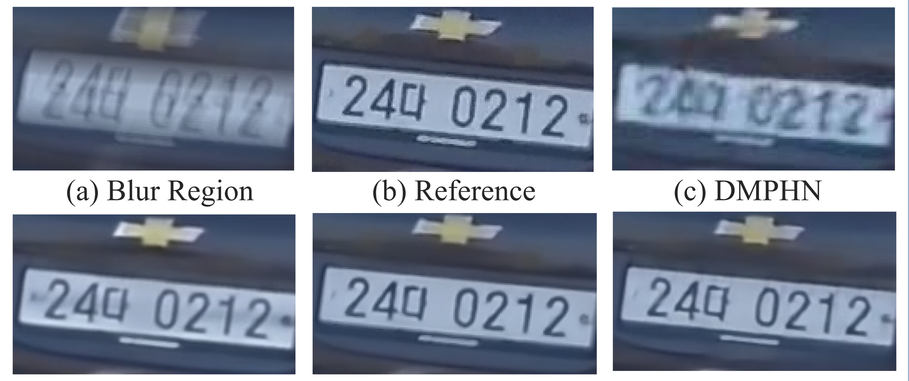


(a) Rainy Image (b) GT (c) MPRNet (d) MAXIM (e) Restormer (f) Ours

Figure 3. Qualitative comparison results on the rain streak removal.

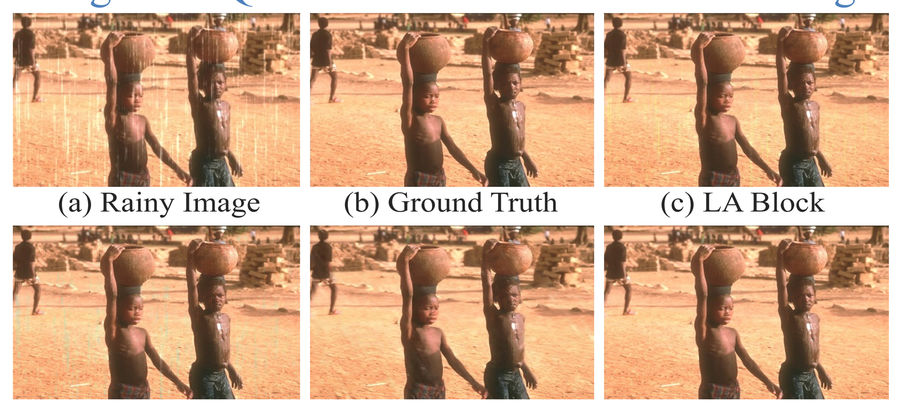
Methods	PSNR	SSIM	FLOPs	Param (M)
MIMO	32.68	0.959	617	16.1
MPRNet	32.66	0.959	777	20.1
Restormer	32.92	0.961	140	26.1
IRNeXt	33.16	0.962	114	13.21
Stripformer	33.08	0.962	170	20.0
Ours	**33.25**	**0.964**	**55**	**4.6**

Table 2. Quantitative results on the GoPro (Deblurring)



(a) Blur Region (b) Reference (c) DMPHN (e) IRNeXt (f) Restormer (g) Ours

Figure 4. Qualitative results on the deblurring



(a) Rainy Image (b) Ground Truth (c) LA Block (d) 1D scan Mamba (e) 2D scan Mamba (f) LA Mamba

Figure 5. The results of different scan mechanism