Local Implicit Wavelet Transformer for Arbitrary-Scale Super-Resolution

Shanghai Key Lab of Medical Image Computing and Computer-assisted Intervention

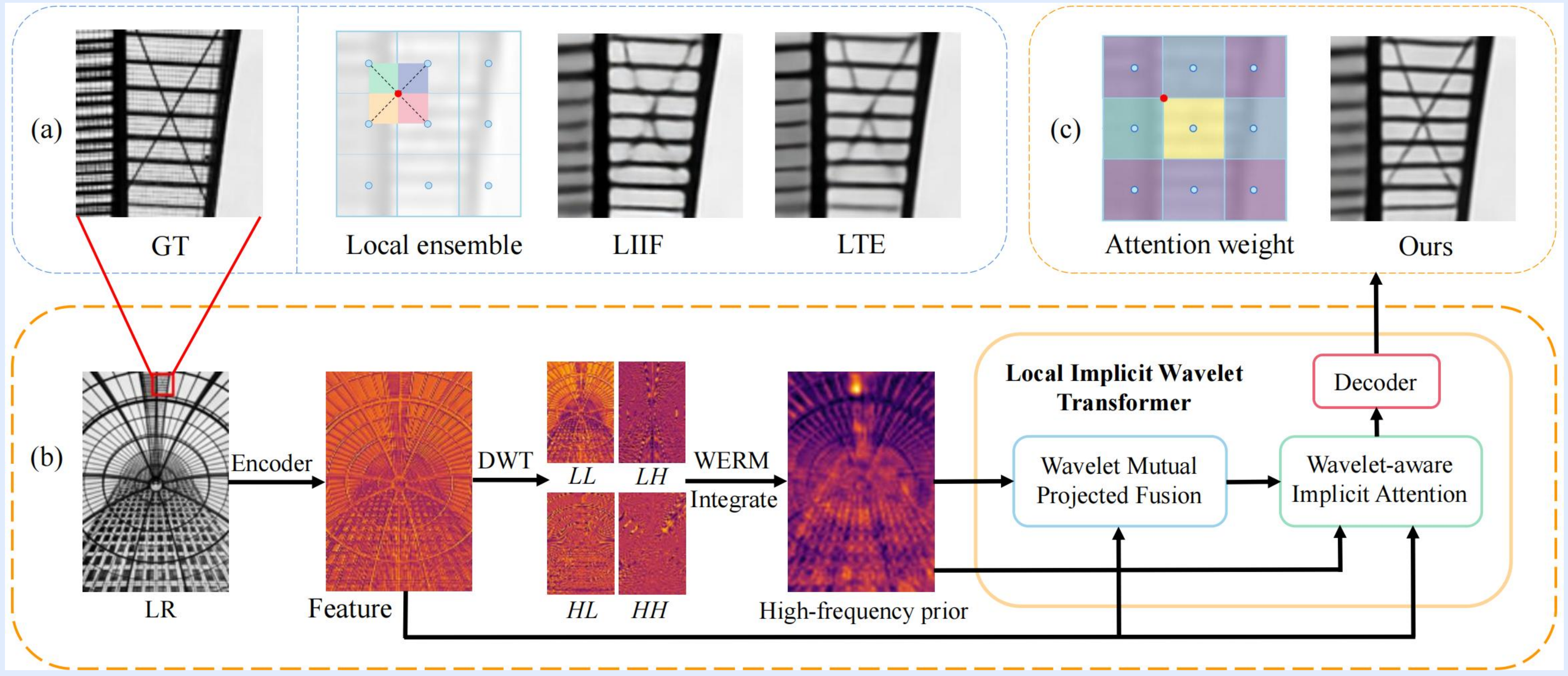
Minghong Duan, Linhao Qu, Shaolei Liu, Manning Wang*

BMVC 2024

Motivation

There are two limitations to existing arbitrary-scale SR methods:

- The coordinate-based local ensemble technique used for querying RGB values fails to consider the relevance of features within local regions.
- The representational capacity of LR features directly obtained from the encoder is limited, overlooking the importance of incorporating high-frequency prior information in images.

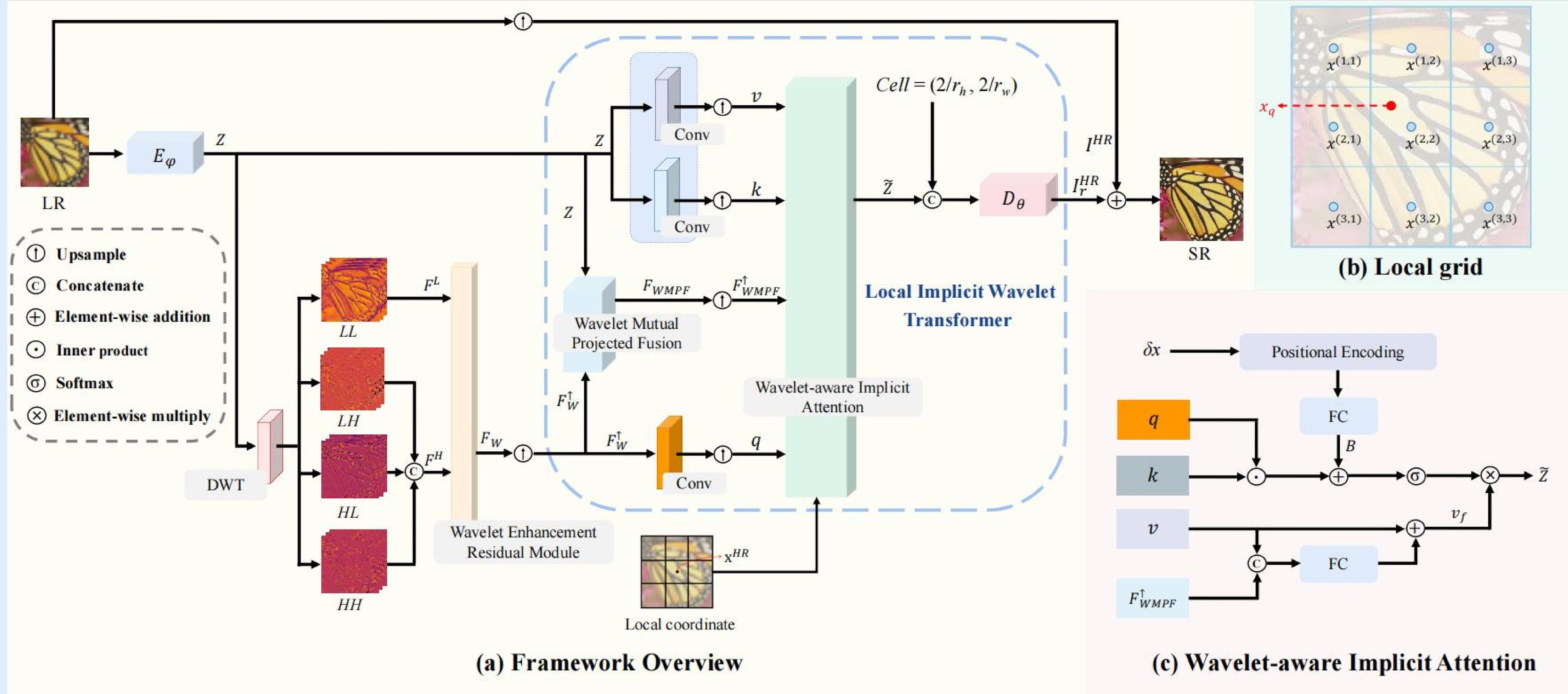


Contributions

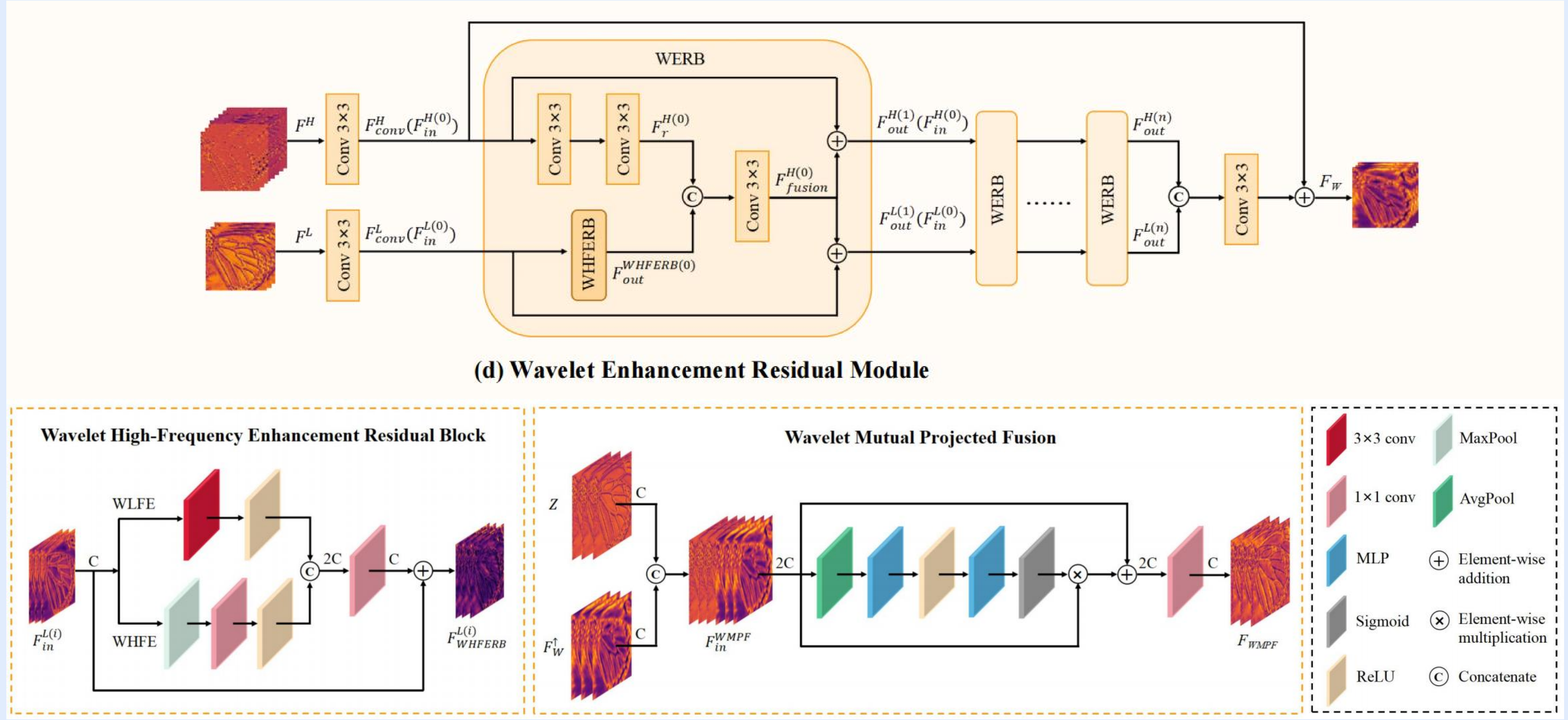
- We introduce the Local Implicit Wavelet Transformer (LIWT), which integrates features from the discrete wavelet transform into local implicit image functions using the Wavelet Mutual Projected Fusion (WMPF) and Wavelet-aware Implicit Attention (WIA), to enhance the learning of high-frequency details.
- We demonstrate that LIWT can be effectively integrated into different encoders to enhance performance, outperforming other arbitrary-scale SR methods.
- We conduct a comprehensive analysis of LIWT. Extensive experimental results demonstrate that the proposed LIWT can produce superior or comparable results on benchmark datasets.

Methods

Framework Overview and Wave-aware Implicit Attention



Wavelet Enhancement Residual Mudule and Wavelet Mutual Projected Fusion



Experiments

Quantitative comparison on DIV2K validation set (PSNR)

Backbone	Methods	×2	×3	×4	×6	×12	×18	×24	×30
-	Bicubic	31.01	28.22	26.66	24.82	22.27	21	20.19	19.59
EDSR [16]	EDSR-baseline [16]	34.55	30.9	28.94	-	-	-	-	-
	EDSR-Meta-SR [12]	34.64	30.93	28.92	26.61	23.55	22.03	21.06	20.37
	EDSR-LIIF [5]	34.67	30.96	29.00	26.75	23.71	22.17	21.18	20.48
	EDSR-UltraSR [27]	34.69	31.02	29.05	26.81	23.75	22.21	21.21	20.51
	EDSR-IPE [18]	34.72	31.01	29.04	26.79	23.75	22.21	21.22	20.51
	EDSR-LTE [14]	34.72	31.02	29.04	26.81	23.78	22.23	21.24	20.53
	EDSR-LIWT(ours)	34.79	31.12	29.15	26.91	23.86	22.31	21.30	20.60
RDN [31]	RDN-baseline [31]	34.94	31.22	29.19	-	-	-	-	-
	RDN-Meta-SR [12]	35.00	31.27	29.25	26.88	23.73	22.18	21.17	20.47
	RDN-LIIF [5]	34.99	31.26	29.27	26.99	23.89	22.34	21.31	20.59
	RDN-UltraSR [27]	35.00	31.30	29.32	27.03	23.73	22.36	21.33	20.61
	RDN-IPE [18]	35.04	31.32	29.32	27.04	23.93	22.38	21.34	20.63
	RDN-LTE [14]	35.04	31.32	29.33	27.04	23.95	22.40	21.36	20.64
	RDN-LIWT(ours)	35.07	31.36	29.39	27.11	24.03	22.47	21.43	20.70
SwinIR [15]	SwinIR-baseline [15]	34.94	31.22	29.19	-	-	-	-	-
	SwinIR-Meta-SR [12]	35.15	31.40	29.33	26.94	23.80	22.26	21.26	20.54
	SwinIR-LIIF [5]	35.17	31.46	29.46	27.15	24.02	22.43	21.40	20.67
	SwinIR-LTE [14]	35.24	31.50	29.51	27.20	24.09	22.50	21.47	20.73
	SwinIR-LIWT(ours)	35.25	31.53	29.55	27.25	24.15	22.56	21.52	20.77

Visual comparison at integer scales and non-integer scales