An Efficient QP Variable Convolutional Neural Network Based In-loop Filter for Intra Coding

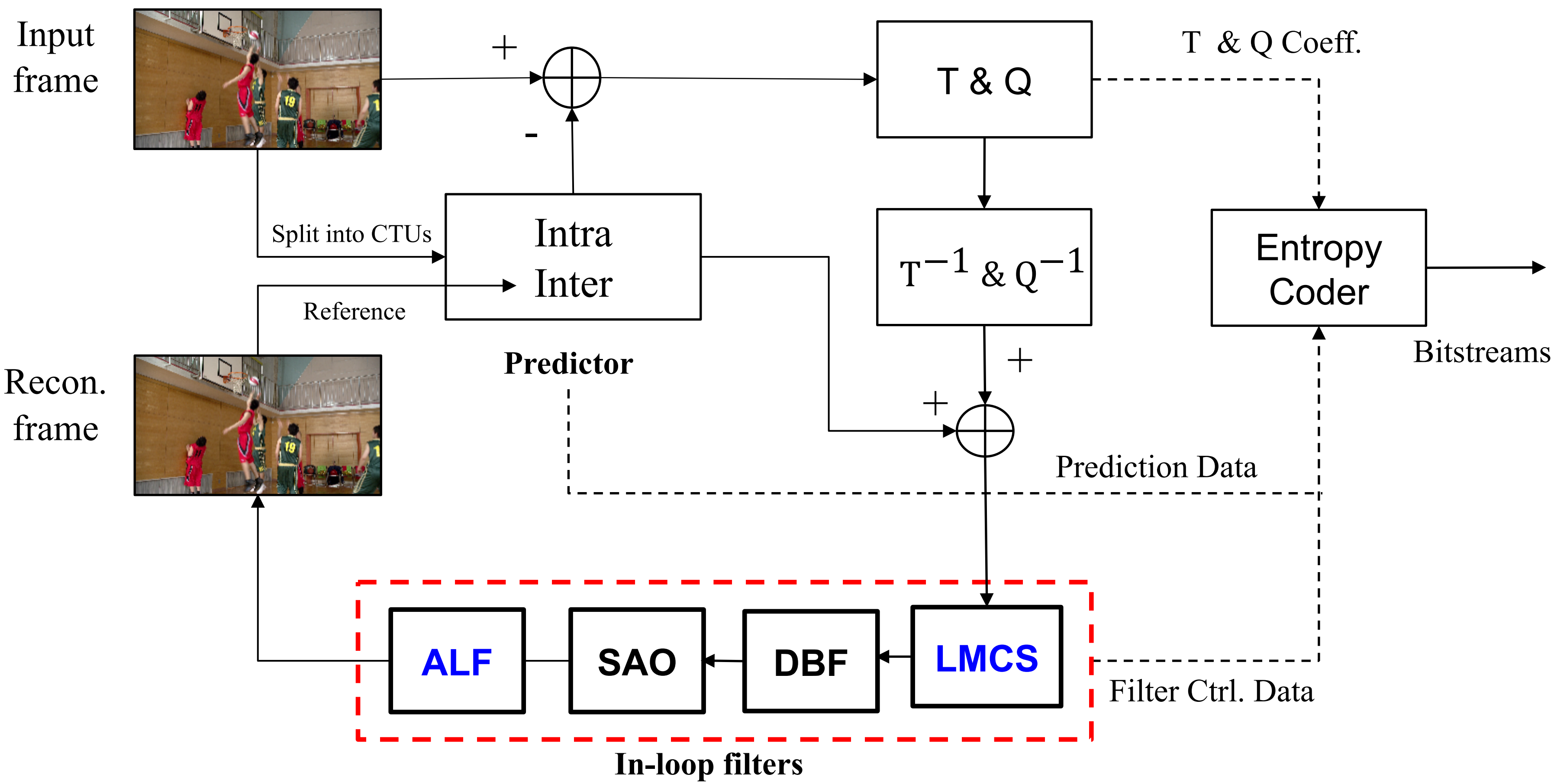
Zhijie Huang(**Speaker**), Xiaopeng Guo, Mingyu Shang, Jie Gao and Jun Sun
Wangxuan Institute of Computer Technology, Peking University
2021.03.24

Outline

- Introduction
- Proposed Method
- Results & Conclusion

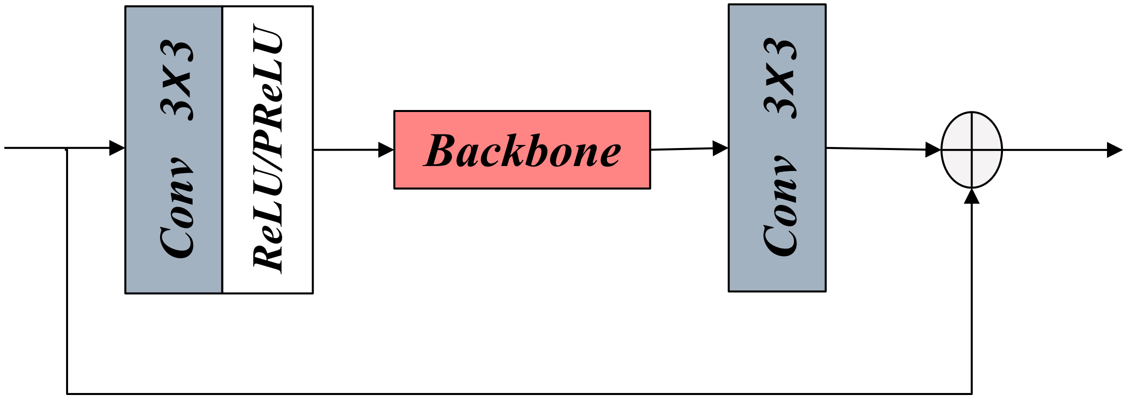
Introduction

- In-loop Filters in Versatile Video Coding (VVC)



Introduction

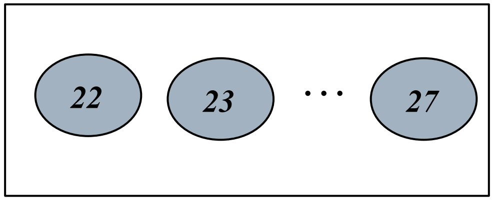
- CNN Based In-loop Filters
  - RRCNN
  - RHCNN
  - DRNLF
  - …

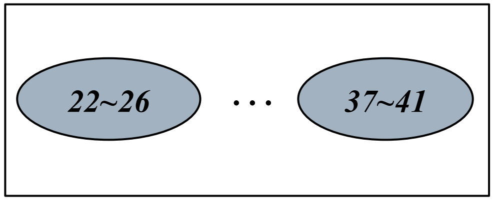


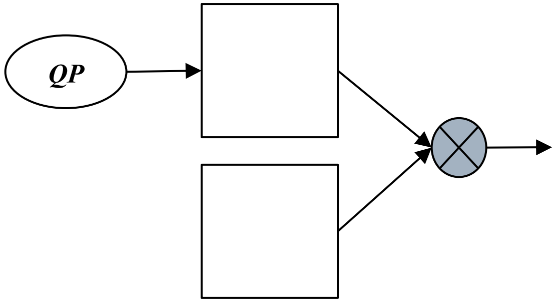
One popular architecture used by many CNN based in-loop filters

Introduction

- Be adaptive to different quantization parameters
  - One model for each QP
  - QP-band method
    - Merged the QPs into several bands, and trained the optimal models for each band.
  - QP Map method
    - Combined QPs as an input and fed them into the CNN training stage by simply padding the scalar QPs into a matrix with the same size of input frames or patches.

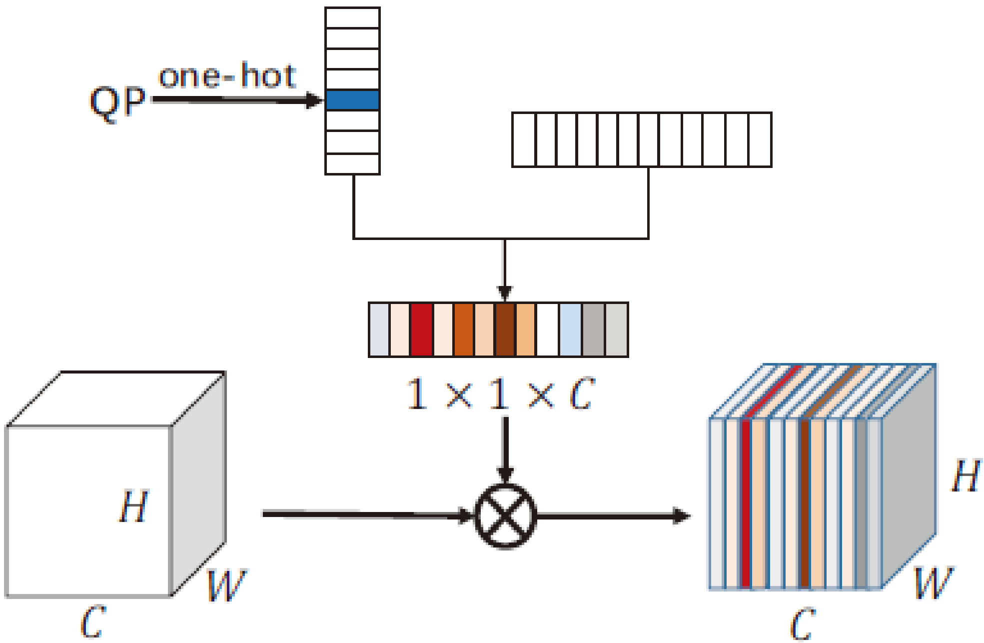






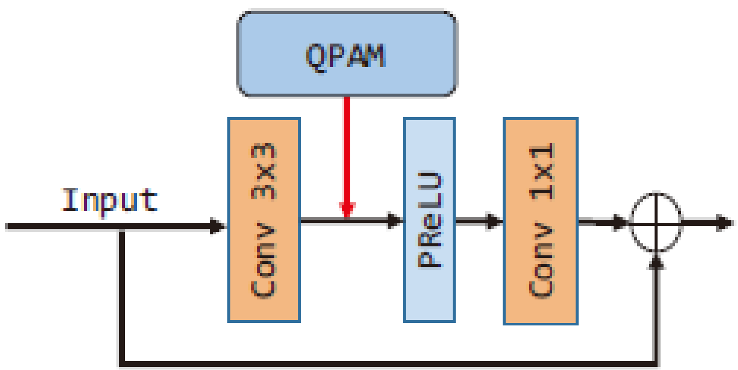
Proposed Method

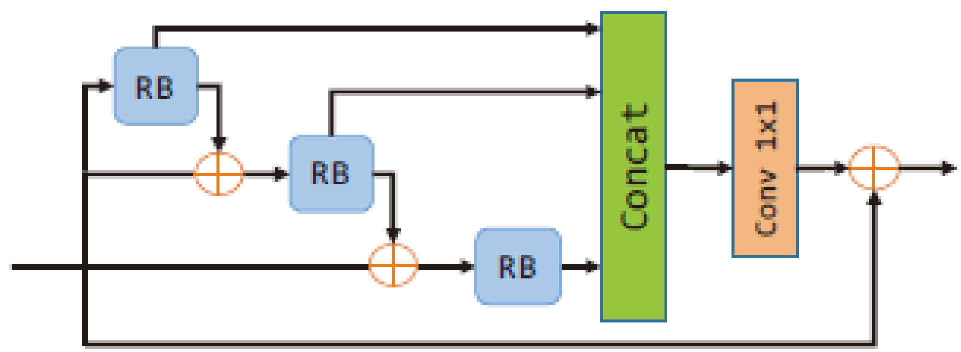
- QP attention module (QPAM)
  - $F' = M \otimes F$
- The QP attention map M
  - $M' = \sigma(Uv_{\Omega}(q))$
  - $M = reshape(M')$
  - $\sigma(x) = log(1 + e^{x}).$



Proposed Method

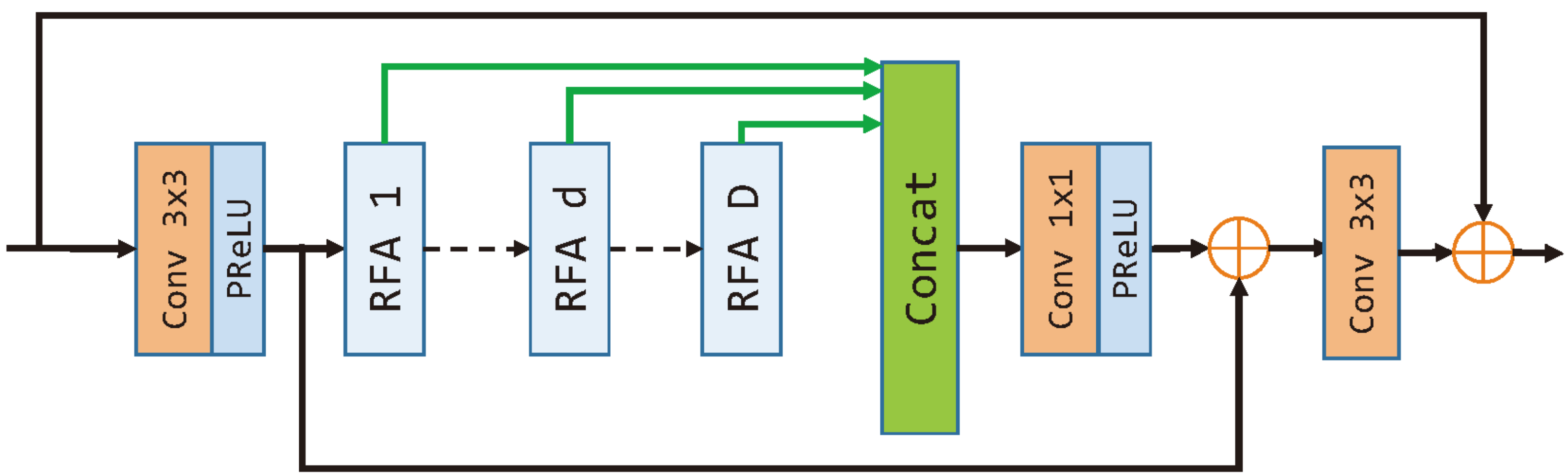
- Residual Block with QPAM
- Residual Feature Aggregation Module





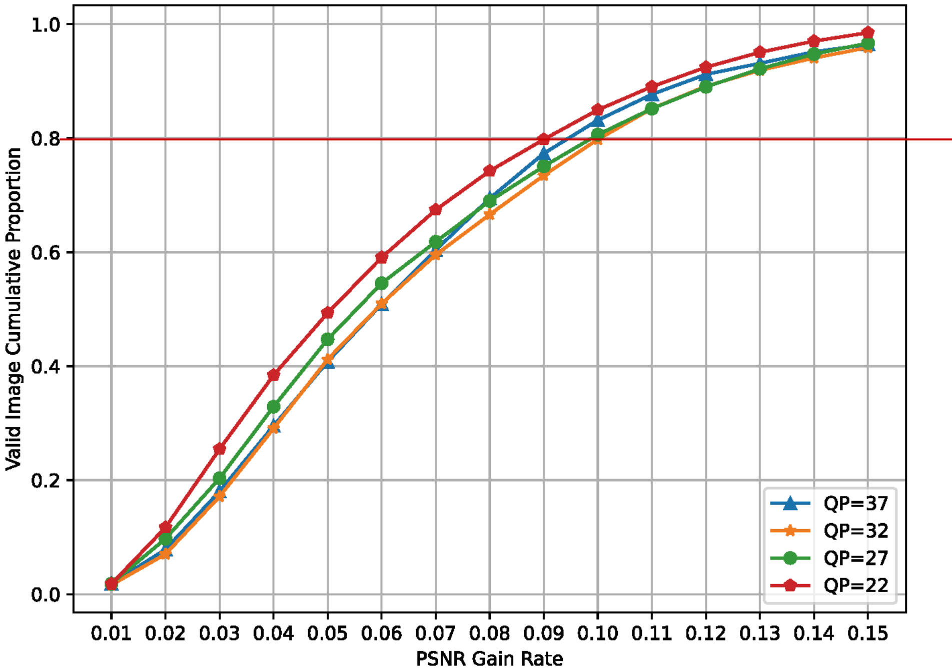
Proposed Method

- The architecture of our QPALF network
  - Also use the popular architecture.
  - Backbone: D cascaded residual feature aggregation modules.



Proposed Method

- PSNR Gain Rate
  - $L_{rec} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{Y} - Y\right|^2$
  - $R = 1 - \frac{L_{rec}}{L_{init}}$
- Loss Function
  - $L = \alpha_q(1 - R)^{\gamma}\, L_{rec}$



Results

- Training Detail
  - Batch Size: 64
  - Epoch: 50
  - Learning rate: 10^{-4}, fine-tune 10^{-5}
- Implementation
  - As an additional tool of in-loop filters between DBF and SAO
  - Incorporated into the VVC reference software VTM6.0
- Test Dataset
  - A1-E recommended by JVET
  - All-intra configuration

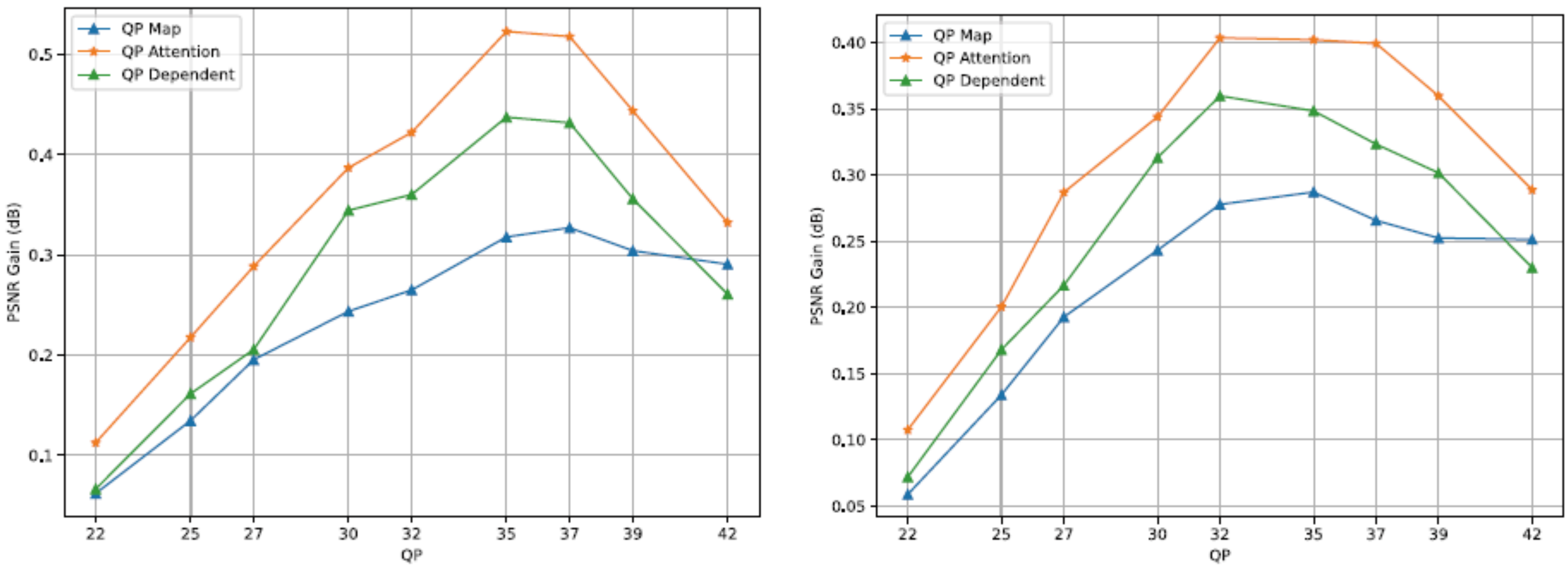
Results

- RD performance

Class	Sequence	BD-Rate(%)				
		RHCNN	DRNLF	QPMLF	QPALF-S	QPALF
A1	Tango2	-0.62	-0.63	-0.78	-0.62	-1.86
	Campfire	-0.82	-1.32	-0.79	-1.42	-2.01
	FoodMarket4	-0.74	-0.09	-0.29	-0.89	-2.20
A2	CatRobot	-1.10	-2.20	-1.97	-2.28	-3.39
	DaylightRoad2	-0.43	0.07	-0.17	1.02	-0.50
	ParkRunning3	-1.01	-1.96	-1.56	-2.04	-2.99
B	RitualDance	-1.88	-4.29	-4.03	-4.85	-6.32
	MarketPlace	-1.34	-2.33	-2.09	-2.64	-3.58
	BasketballDrive	-0.73	-1.63	-1.05	-1.82	-2.84
	BQTerrace	-0.64	-1.38	-1.02	-1.56	-2.06
	Cactus	-0.84	-2.07	-1.71	-1.69	-1.78
C	BasketballDrill	-2.29	-5.43	-4.50	-5.76	-7.48
	BQMall	-1.93	-4.31	-3.73	-4.58	-5.49
	PartyScene	-1.22	-3.01	-2.57	-3.19	-3.62
	RaceHorsesC	-0.81	-1.75	-1.39	-1.81	-2.31
D	BasketballPass	-2.10	-5.24	-4.42	-5.67	-6.76
	BlowingBubbles	-1.59	-3.64	-3.19	-3.87	-4.45
	BQSquare	-1.93	-5.12	-4.40	-5.28	-6.20
	RaceHorses	-2.03	-4.54	-4.21	-4.67	-5.40
E	FourPeople	-2.05	-4.73	-4.08	-5.09	-6.49
	Johnny	-1.63	-3.90	-3.17	-4.12	-5.72
	KristenAndSara	-1.59	-3.95	-3.31	-4.24	-5.31
Average All		**-1.54**	**-2.88**	**-2.47**	**-3.05**	**-4.03**

Results

- PSNR gain on multi-QPs



(a) FourPeople (b) BQMall

Results

Subjective Comparison

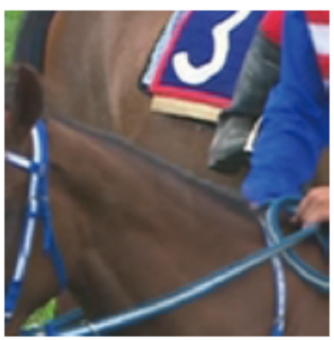
Ground Truth
(PSNR,SSIM)

VTM
(26.16, 0.9045)

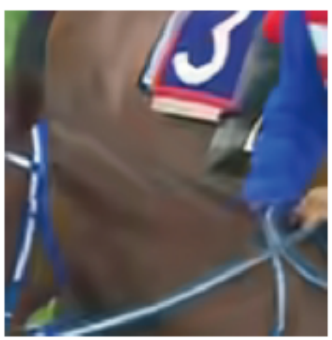
RHCNN
(26.29, 0.9073)

DRNLF
(26.56, 0.9104)

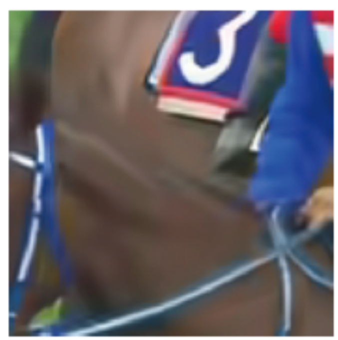
Ours
(**26.66, 0.9120**)



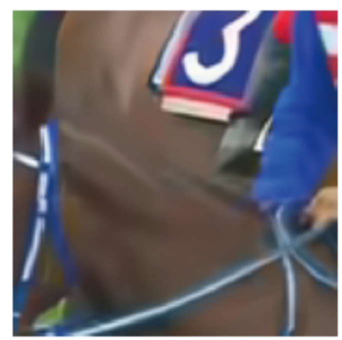
Ground Truth
(PSNR,SSIM)



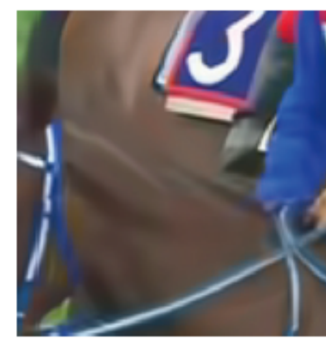
VTM
(28.15, 0.8643)



RHCNN
(28.24, 0.8666)



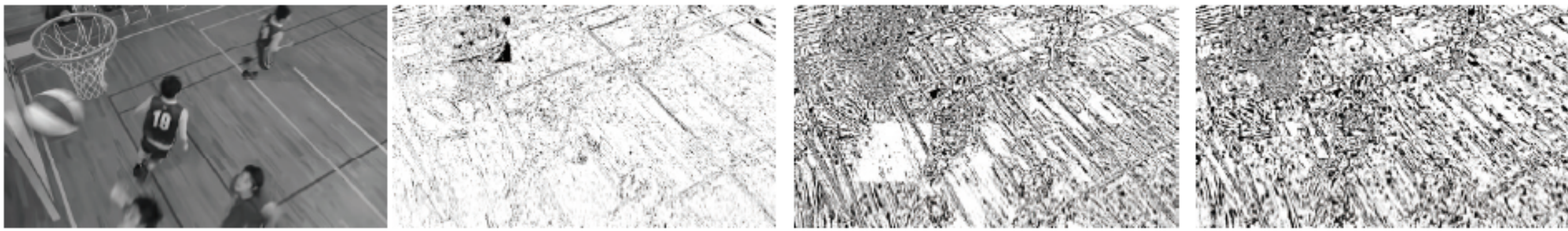
DRNLF
(28.32, 0.8682)



Ours
(**28.42, 0.8714**)

Results

Subjective Comparison



(a) VTM (b) RHCNN (c) DRNLF (d) Ours

Complexity Comparison

Method	ΔET	ΔDT	#Params
RHCNN	5.43%	10695.9%	$6.79M \times 4$
DRN	3.29%	7808.0%	$671.30k \times 4$
QPMLF	4.41%	8232.2%	$838.98k \times 1$
QPALF	4.70%	8428.6%	$905.22k \times 1$

Conclusion

- Contribution
  - We propose an efficient QP attention module
  - We design a network architecture that is equipped with controllability for different QPs.
  - We introduce a focal MSE to train a more robust model
- Future Work
  - Optimize our QP attention module
  - Extend our model to inter mode

Thank You!

Q&A

zhijiehuang@pku.edu.cn