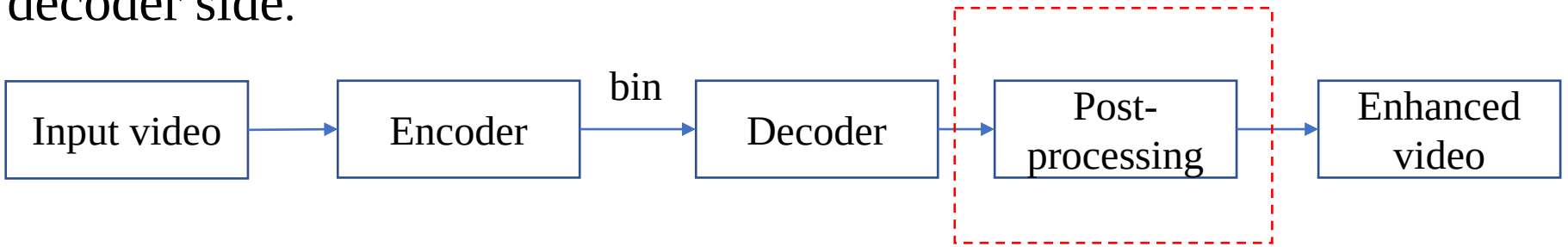
A Low-complexity Neural Network for Compressed Video Post-processing in HEVC

Introduction

Post-processing

Video post-processing is a method to improve the quality of reconstructed frames at the decoder side.



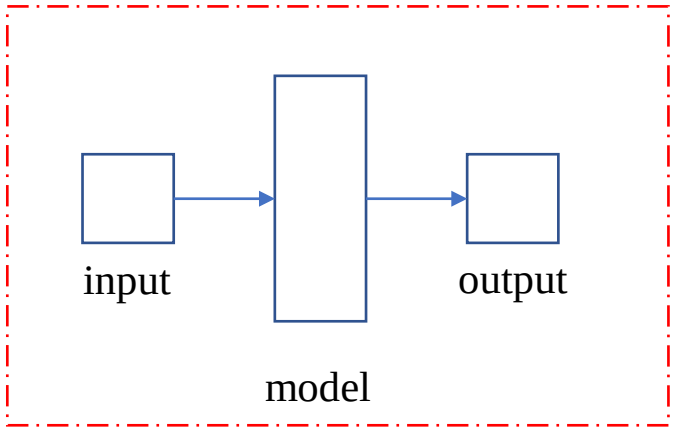
Video enhancement process

The input video forms a bitstream through the encoder, and then the decoder decodes the bitstream into video. After that, the reconstructed video quality is improved by post-processing method.

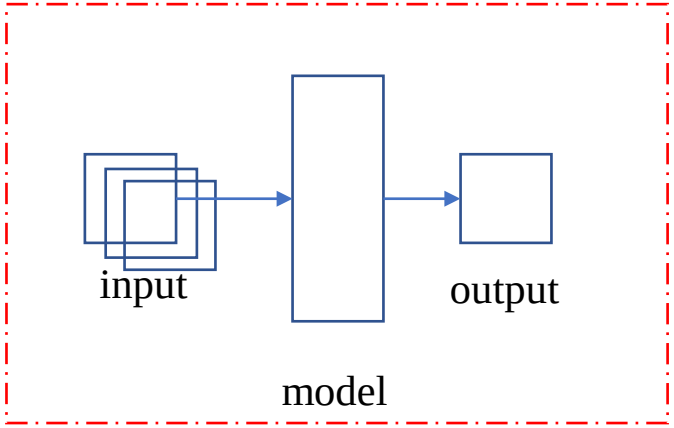
Introduction

Learning-based post-processing method

- Single-frame post-processing
- Multi-frame post-processing



Single-frame



Multi-frame

Motivation

The existing methods have the following characteristics:

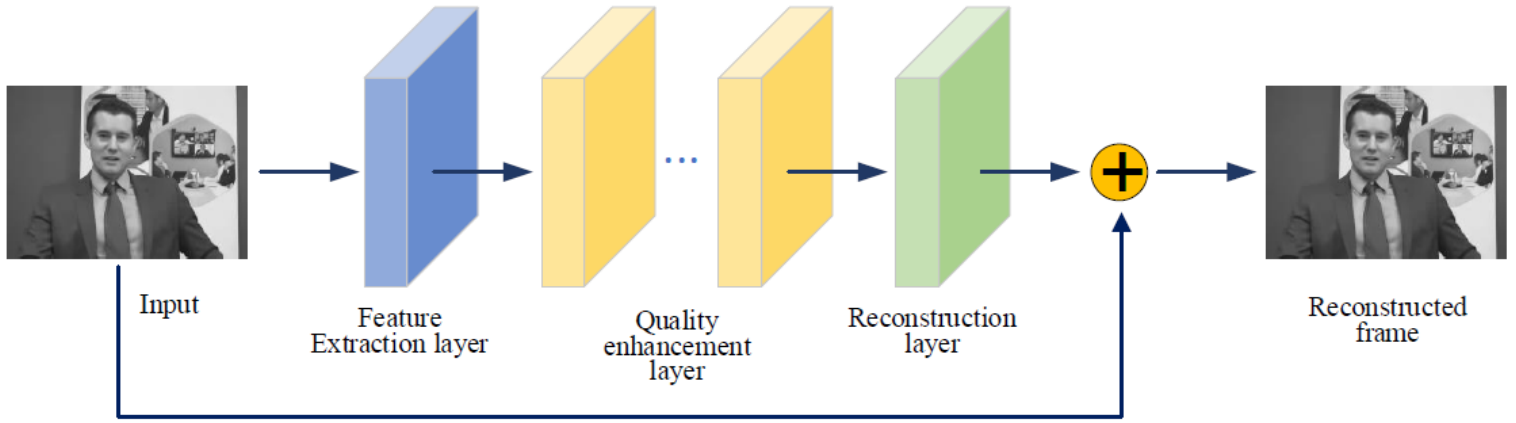
- Heavy computational burden.
- Too much memory usage.

To tackle these problems, we design a low-complexity neural network

- Depth-wise separable convolution is introduced to reduce the computational complexity and improve the processing speed.
- Max-pooling is introduced to reduce memory usage.
- Shortcut is used to reuse features of different scales.

Proposed Method

Architecture

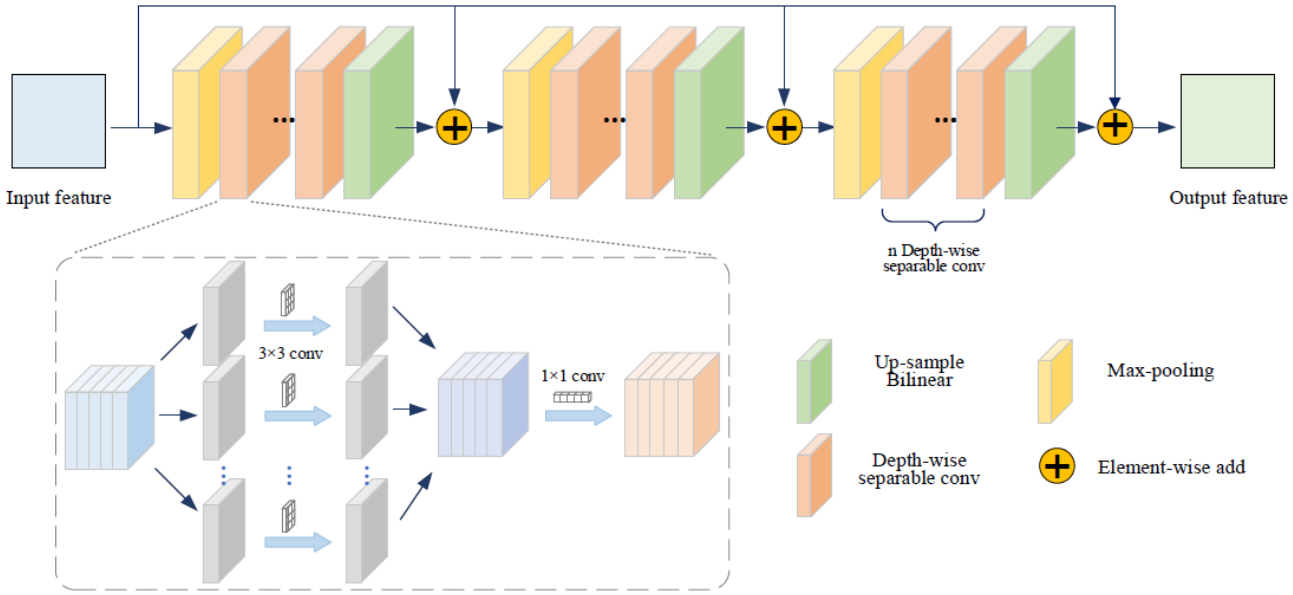


Overview of the proposed model

The model includes three parts named feature extraction layer, quality enhancement layer and reconstruction layer.

Proposed Method

Architecture



Quality enhancement layer

Depth-wise separable convolution is used to reduce the complexity of the model.

Experiment Result

FLOPs and parameters

Table 3: Parameters and FLOPs between different models

	ARCNN[19]	QECNN-I[13]	DCAD[12]	Ours
Params(K)	106.56	454.66	296.64	**61.21**
FLOPs(G)	82.86	353.54	230.67	**33.83**

Suppose the resolution of the input frame is 1280x720.

Experiment Result

Quality Results

Resolution	Sequence name	ARCNN[19]	QECNN-I[13]	DCAD[12]	Ours
ClassB	BasketballDrive	0.101/11.63	0.138/4.98	0.244/8.32	**0.266**/**12.96**
1920x1080	BQTerrace	0.161/11.63	0.246/4.98	**0.416**/8.34	0.414/**13.06**
	Cactus	0.107/11.58	0.141/4.96	0.237/8.31	**0.256**/**13.11**
	Kimono	0.113/11.61	0.134/4.98	**0.196**/8.32	0.187/**13.11**
	ParkScene	0.082/11.54	0.097/4.94	**0.165**/8.29	0.161/**13.05**
1920x1080	Average	0.113/11.60	0.151/4.97	0.252/8.31	**0.257**/**13.06**
ClassC	BasketballDrill	0.188/58.29	0.291/25.20	**0.518**/41.28	0.469/**63.61**
832x480	BQMall	0.168/57.66	0.196/25.25	0.323/41.36	**0.333**/**63.97**
	PartyScene	0.121/58.37	0.167/25.18	**0.273**/41.21	0.267/**63.75**
832x480	Average	0.159/58.11	0.218/25.21	**0.371**/41.28	0.356/**63.77**
ClassD	BasketballPass	0.121/**203.5**	0.193/92.36	**0.383**/145.1	0.380/160.4
416x240	BlowingBubbles	0.090/**199.6**	0.122/91.15	0.209/146.3	**0.225**/160.7
	BQSquare	0.1419/**199.2**	0.314/90.69	**0.562**/145.7	0.539/159.7
	RaceHorses	0.2716/**198.9**	0.378/91.04	**0.585**/141.5	0.560/157.1
416x240	Average	0.156/**200.4**	0.252/91.31	**0.435**/144.9	0.426/159.5
ClassE	FourPeople	0.271/25.86	0.332/11.08	0.520/18.44	**0.537**/**28.97**
1280x720	Johnny	0.229/26.14	0.273/11.11	0.396/18.55	**0.438**/**28.89**
	KristenAndSara	0.302/26.04	0.395/11.10	0.532/18.50	**0.583**/**28.00**
	vidyo1	0.265/26.09	0.340/11.13	**0.522**/18.46	0.497/**28.70**
1280x720	Average	0.267/26.03	0.335/11.10	0.492/18.49	**0.514**/**28.64**

ΔPSNR denotes the difference between output and HEVC baseline. FPS denotes the processing speed of the model. Higher values indicate better performance. The black font is the best.

Compared with other methods, our method achieves a compromise effect.

Thanks !