深度学习与视频编码

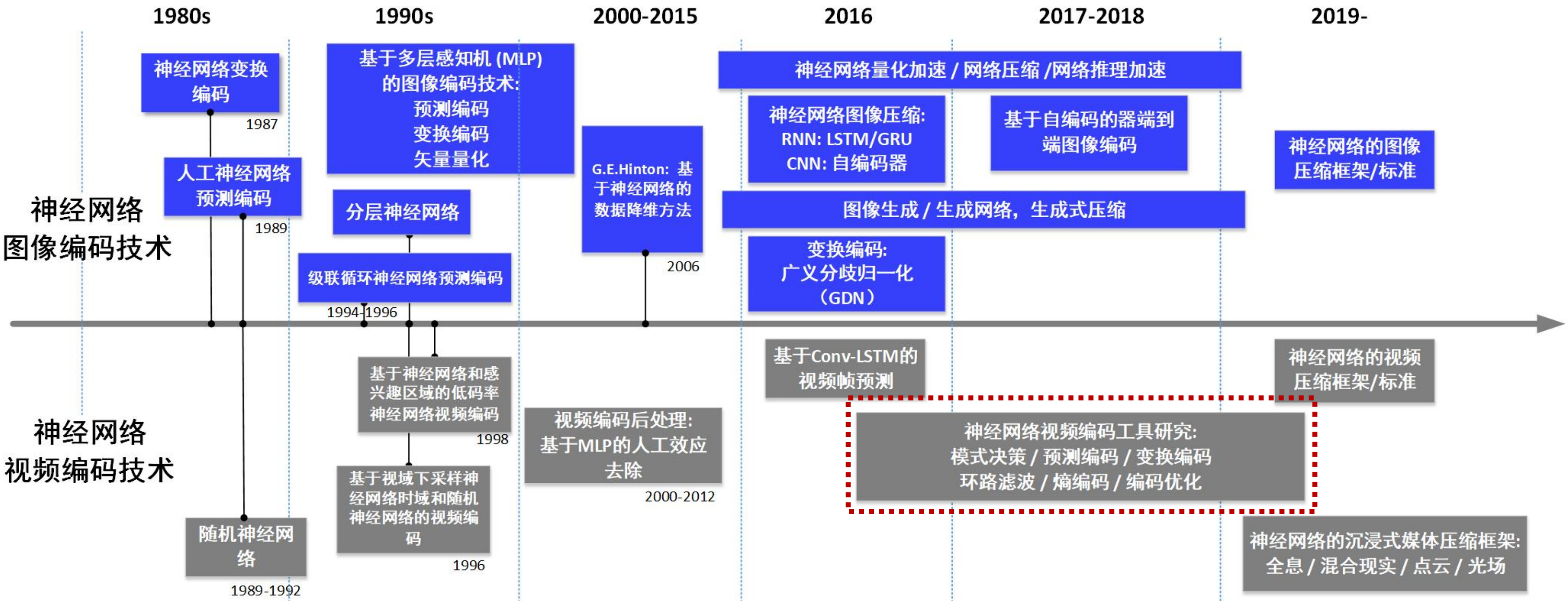
王苫社
北京大学

目录

目录

- 概述
- 神经网络视频编码历史
- 基于深度学习的视频编码进展
- 展望

神经网络视频编码简史



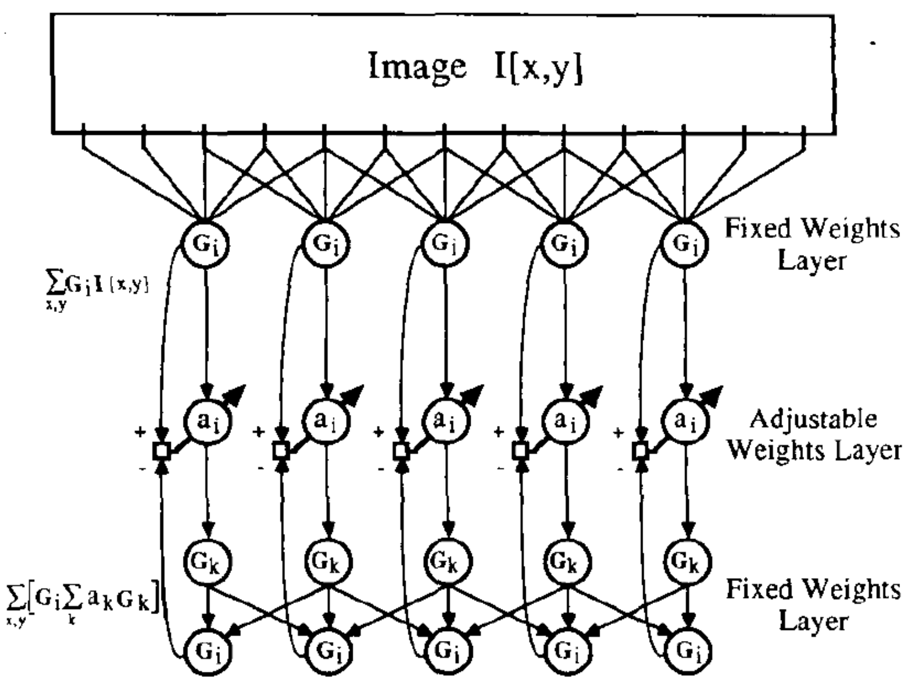
1980s
1990s
2000-2015
2016
2017-2018
2019-
神经网络
图像编码技术
神经网络
视频编码技术
神经网络变换编码
1987
人工神经网络预测编码
1989
基于多层感知机 (MLP) 的图像编码技术:
预测编码
变换编码
矢量量化
分层神经网络
级联循环神经网络预测编码
1994-1996
G.E.Hinton：基于神经网络的数据降维方法
2006
神经网络量化加速 / 网络压缩 /网络推理加速
神经网络图像压缩:
RNN: LSTM/GRU
CNN: 自编码器
基于自编码的器端到端图像编码
神经网络的图像压缩框架/标准
图像生成 / 生成网络，生成式压缩
变换编码:
广义分歧归一化
（GDN）
基于Conv-LSTM的视频帧预测
神经网络的视频压缩框架/标准
基于神经网络和感兴趣区域的低码率神经网络视频编码
1998
基于视域下采样神经网络时域和随机神经网络的视频编码
1996
随机神经网络
1989-1992
视频编码后处理:
基于MLP的人工效应去除
2000-2012
神经网络视频编码工具研究:
模式决策 / 预测编码 / 变换编码
环路滤波 / 熵编码 / 编码优化
神经网络的沉浸式媒体压缩框架:
全息 / 混合现实 / 点云 / 光场

目录

- 概述
- 神经网络视频编码历史
- 基于深度学习的视频编码进展
- 展望

起源

- 基于神经网络的编码技术源自上世纪八十年代.



三层人工神经网络用于图像变换编码. [1, 2]

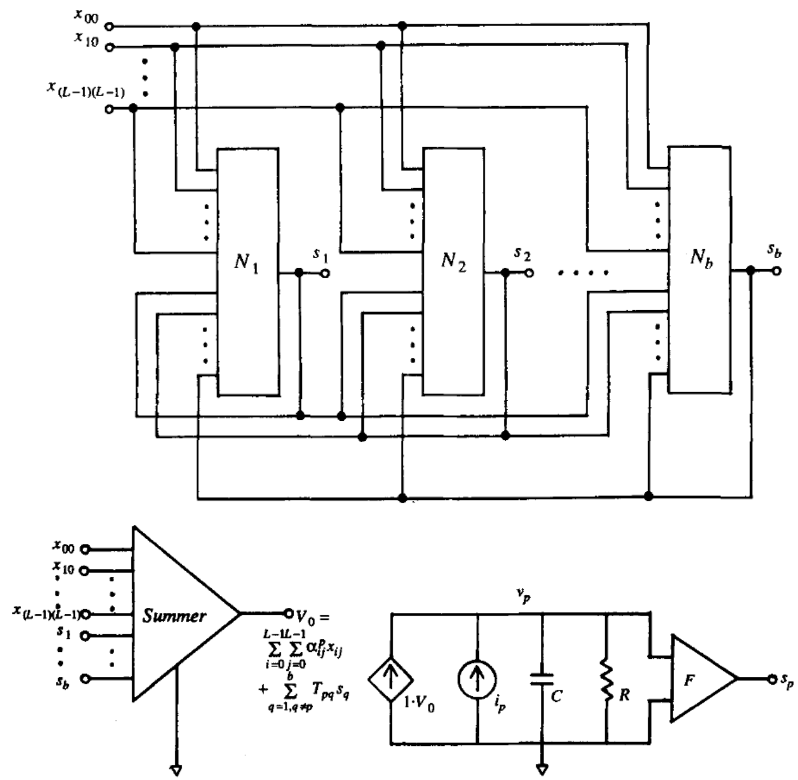
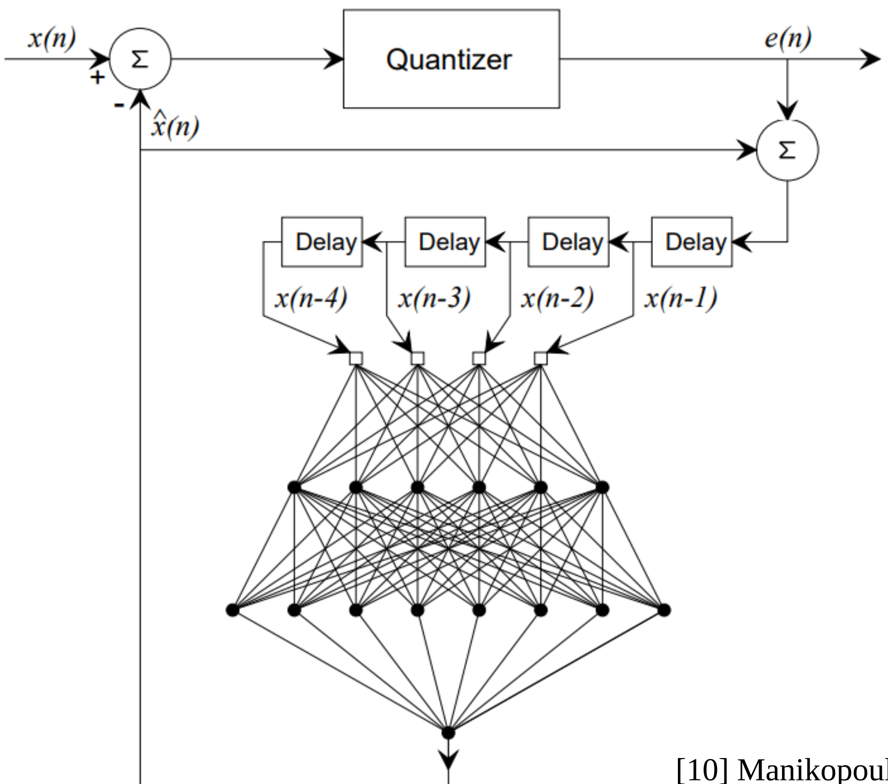


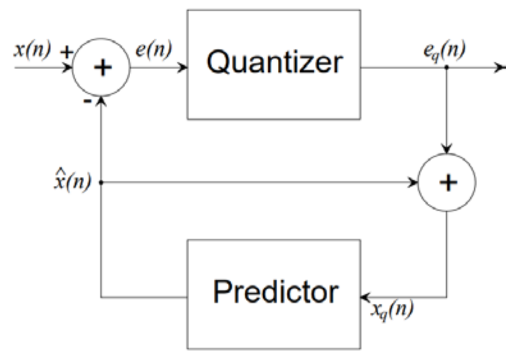
Figure 4. Schematic diagram of a neuron

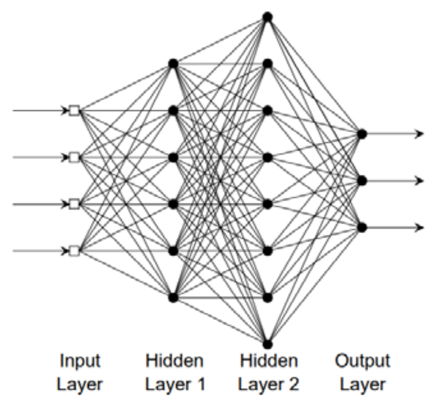
神经网络硬件电路实现 [3]

上世纪九十年代初

- 基于多层感知机的图像编码
  - DPCM using a multilayer perceptron network
  - Predictive Coding







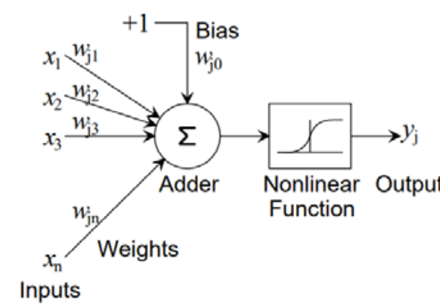
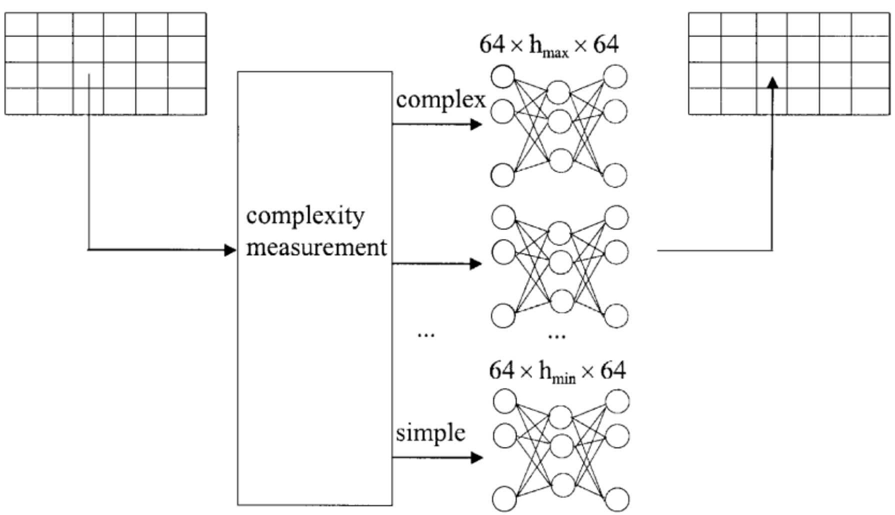


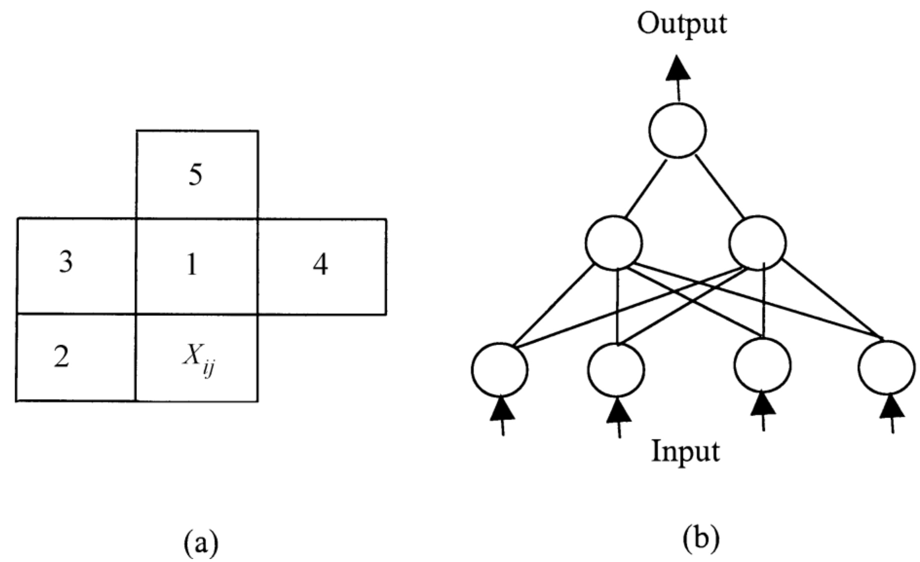
Fig. 3. Model of a neuron.

[10] Manikopoulos C N. Neural network approach to DPCM system design for image coding[J]. IEEE Proceedings I (Communications, Speech and Vision), 1992, 139(5): 501-507.

九十年代中期

- 自适应预测编码 [11-18]
  - Complexity analysis & Entropy coding



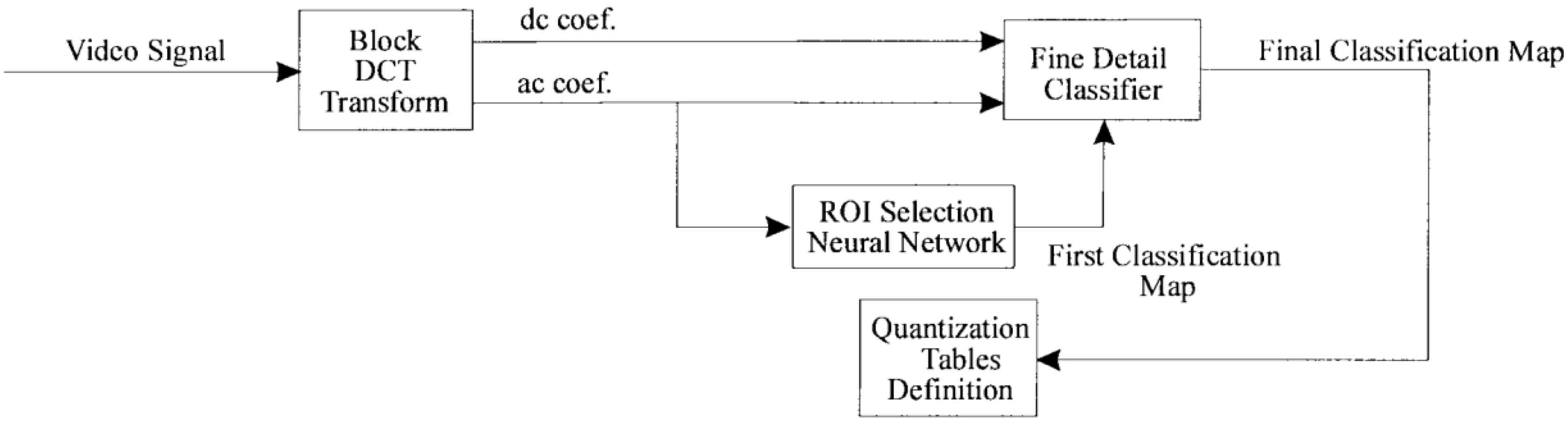


将图像划分为小块进行编码. [11]

利用空域临近像素作为辅助预测. [12, 13]

2000年左右

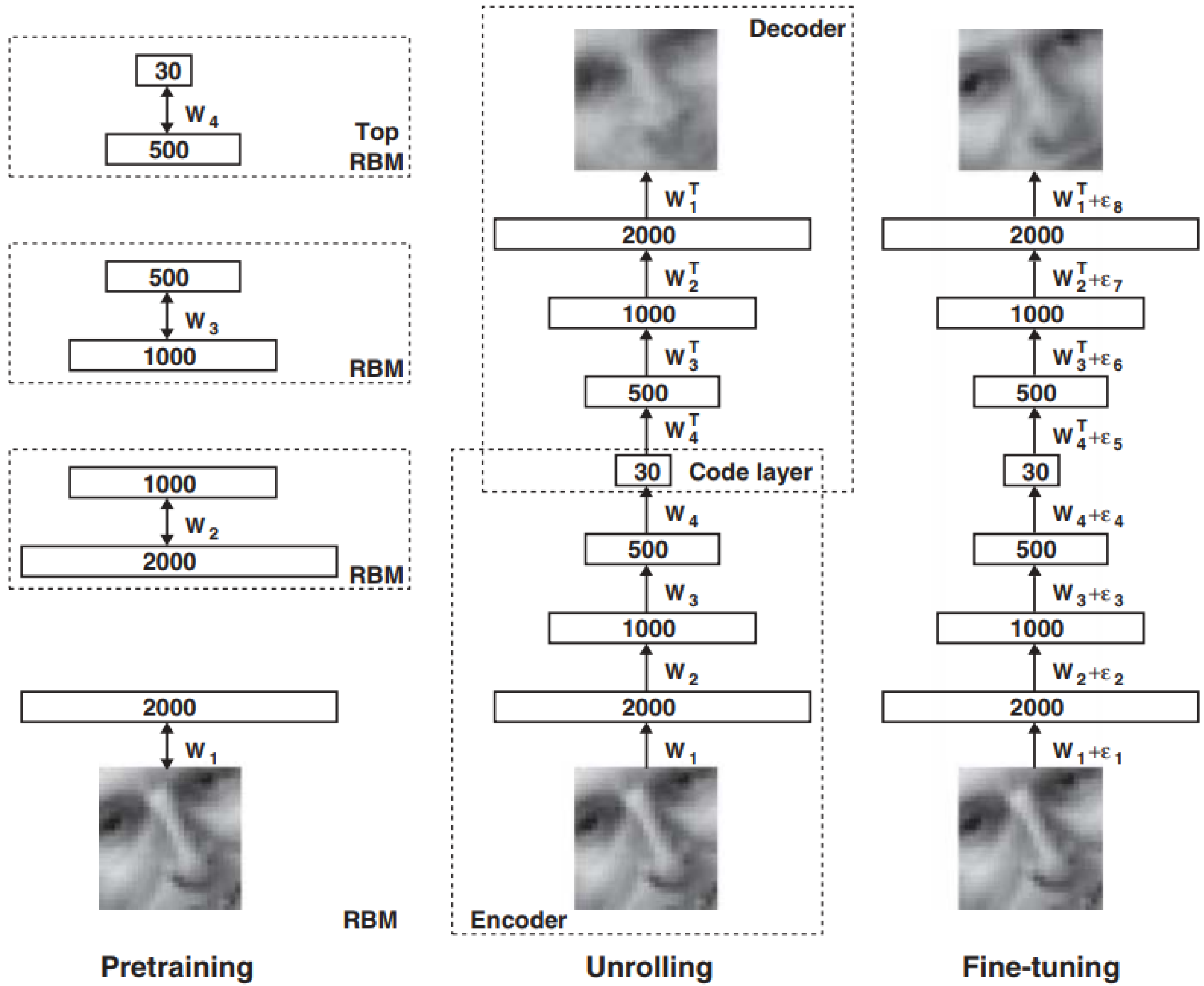
- 两个主要的问题[19-28]
  - 1. 端到端的多层感知机编码
  - 2. 由图像扩展到视频编码



In 2006

- 自编码器开启深度学习时代 by G.E. Hinton

[29] Geoffrey E.et al. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786): 504-507.

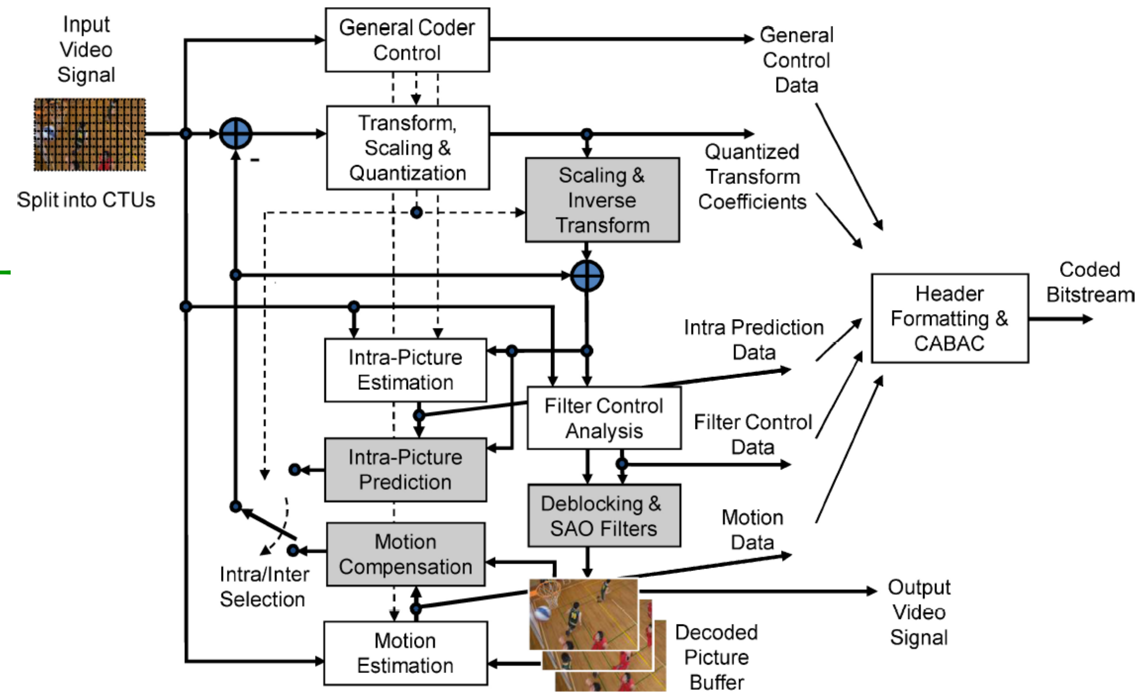


目录

- 概述
- 神经网络视频编码历史
- 基于深度学习的视频编码进展
  - 预测增强
  - 环路滤波
  - 深度学习与视频编码标准
- 展望

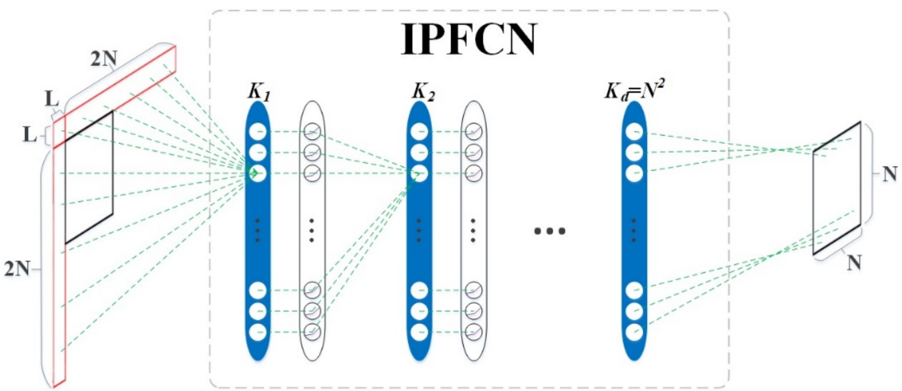
Outline

- 深度学习与视频编码
  - 帧内预测
  - 帧间编码
    - 分像素插值
    - 预测增强
    - 参考帧质量提升
  - 环路滤波
  - 模式决策
    - 编码优化



帧内预测 (1)

- 数据驱动的帧内预测方法
  - 基于全卷积网络
  - 全连接网络
    - 单一模型：IPFCN-S (IPFCN-S-L：将网络参数减半)
    - 双模型（为DC和planar训练专门模型）：IPFCN-D (IPFCN-D-L：将网络参数减半)

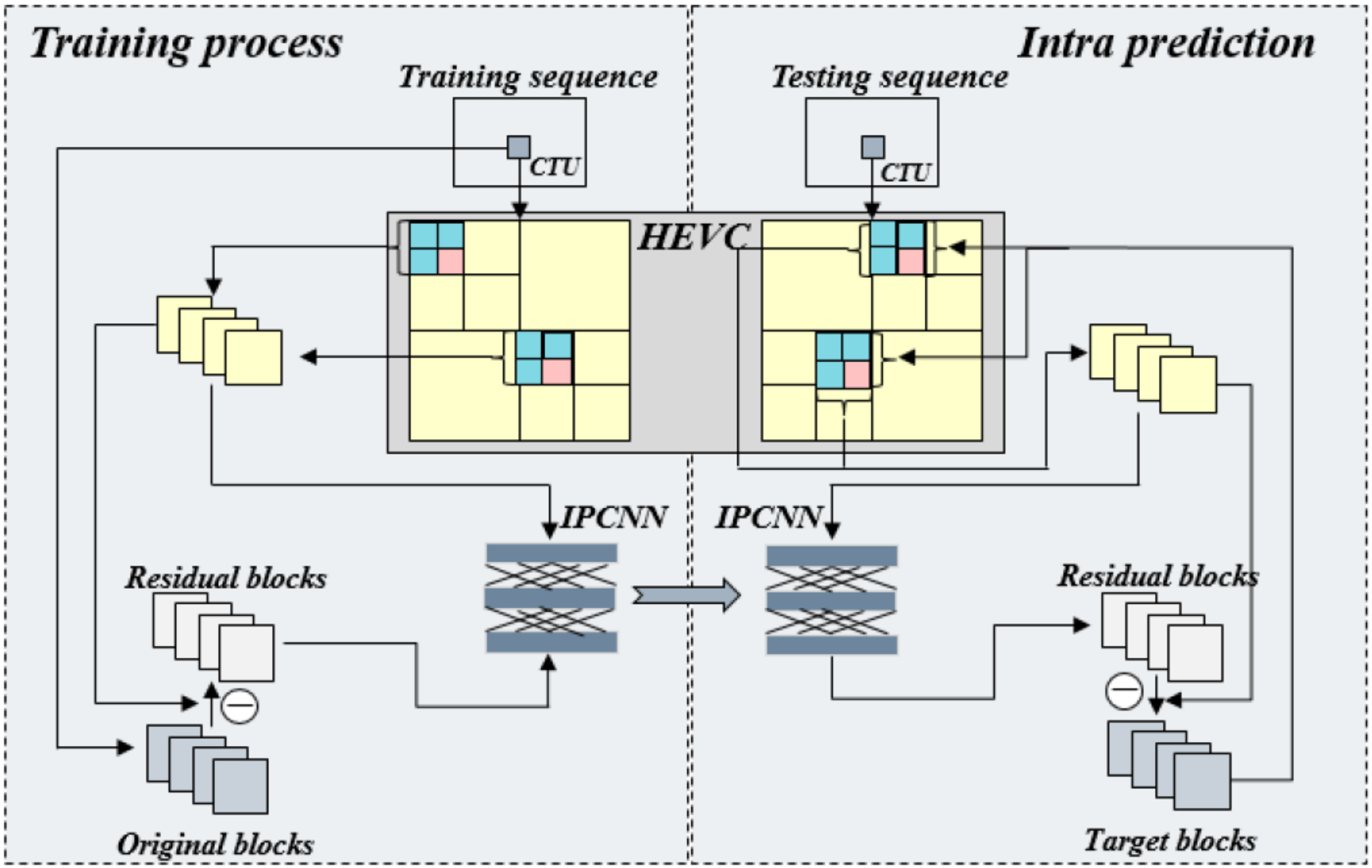


Sequences	IPFCN vs. HM-16.9			
	IPFCN-S	IPFCN-D	IPFCN-S-L	IPFCN-D-L
Class A	-3.8 %	-4.4 %	-3.0%	-3.7%
Class B	-2.8 %	-3.2 %	-2.2%	-2.8%
Class C	-1.9 %	-2.1 %	-1.6%	-1.9%
Class D	-1.7 %	-1.8 %	-1.4%	-1.7%
Class E	-3.9 %	-4.5 %	-3.0%	-3.5%
Overall	-2.6 %	-3.0 %	-2.1%	-2.5%
Encode Time	4930%	13052%	285%	483%
Decode Time	26572%	28927%	923%	1141%

Li, J., Li, B., Xu, J., Xiong, R., Gao, W. Fully connected network-based intra prediction for image coding. IEEE Transactions on Image Processing, 27(7), 3236-3247, 2018.

帧内预测 (2)

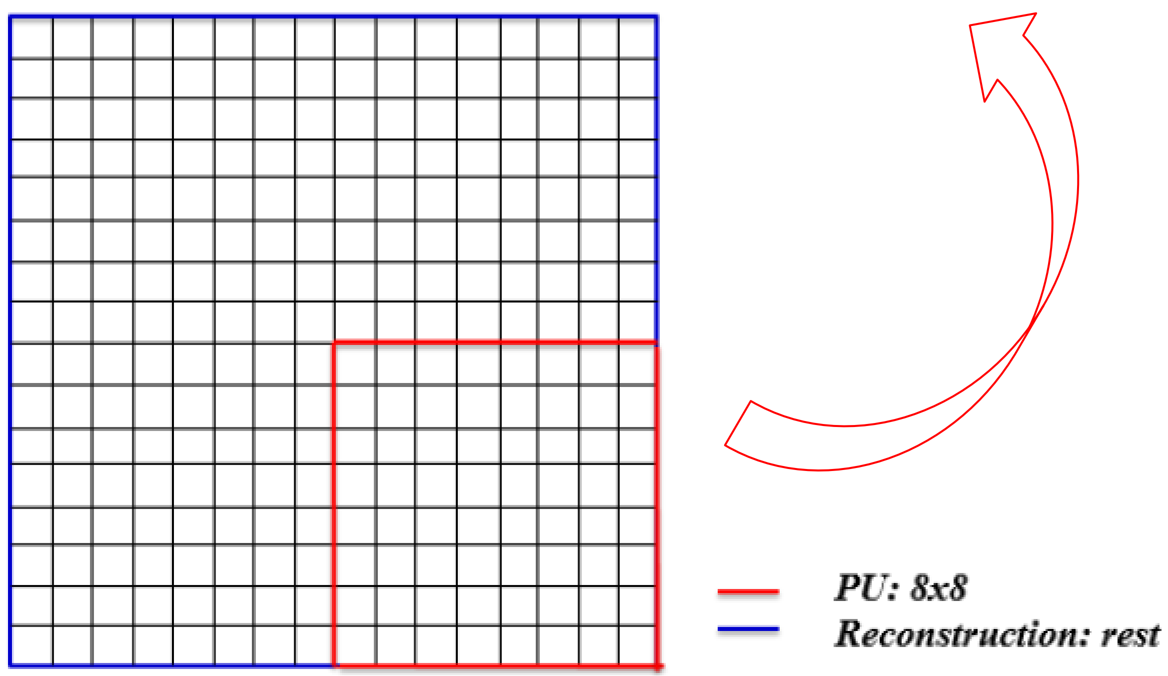
- Design Resnet for Intra 8x8 PU



WX Cui, T Zhang, SP Zhang, F Jiang, WM Zuo, DB Zhao, Convolutional Neural Networks Based Intra Prediction for HEVC, DCC 2017

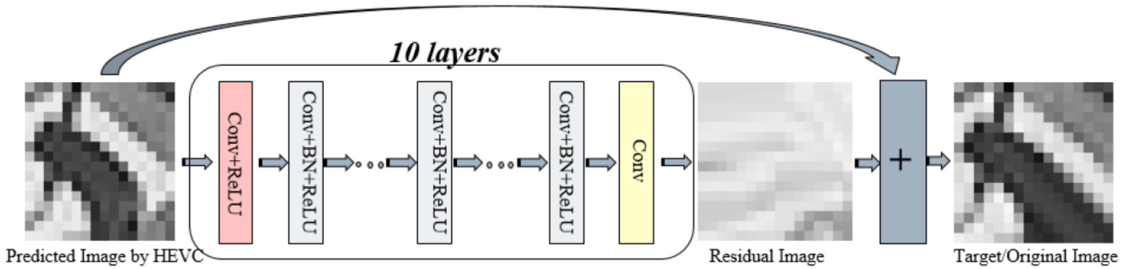
帧内预测(2)

- 预测增强
  - 训练数据生成
  - 当前PU通过HEVC得到最优intra mode



帧内预测(2)

- 残差学习

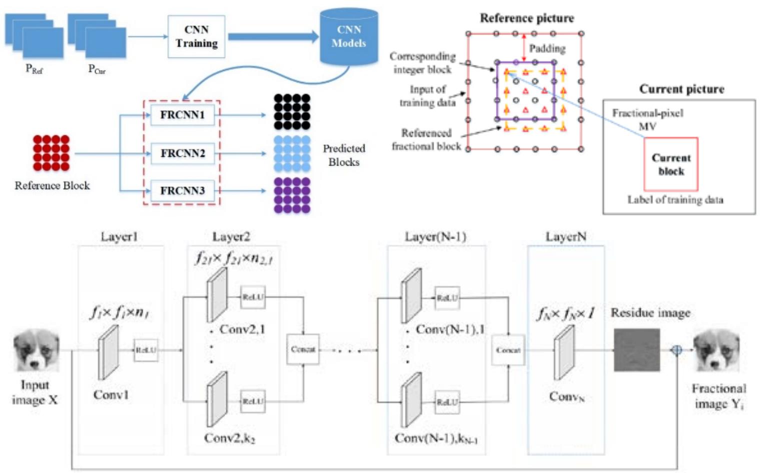


- QP: 22 27 32 37, HEVC Intra, DL Platform: Matcovnet

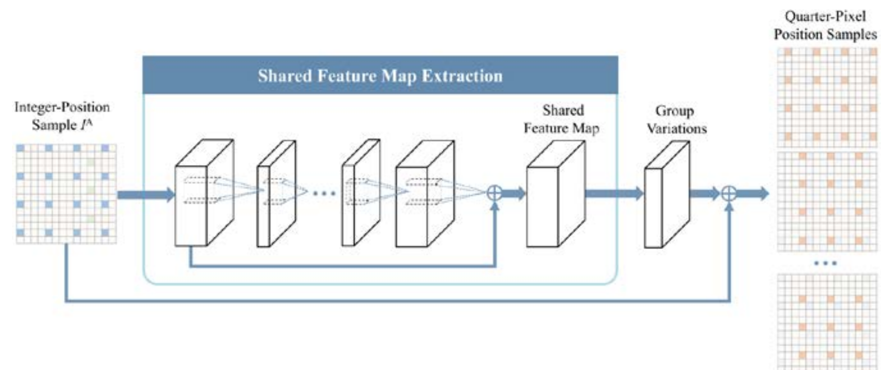
Sequences	BD-rate	Sequences	BD-rate
Traffic	-0.9%	PartyScene	-0.5%
PeopleOnStreet	-1.2%	RaceHorses	-0.7%
Kimono	-0.2%	BasketballPass	-0.4%
ParkScene	-0.8%	BQSquare	-0.1%
Cactus	-0.8%	BlowingBubbles	-0.7%
BasketballDrive	-0.6%	RaceHorses	-0.7%
BQTerrace	-0.8%	FourPeople	-0.3%
BasketballDrill	-0.5%	Johnny	-1.0%
BQMall	-0.6%	KristenAndSara	-0.8%
All average	-0.70%		

基于深度学习的分像素插值

- 针对1/2像素设计神经网络
  - 帧间预测：分像素插值+图像超分辨技术



Class	Sequence	BD-rate of LDP			BD-rate of LDB			BD-rate of RA		
		Y (%)	U (%)	V (%)	Y (%)	U (%)	V (%)	Y (%)	U (%)	V (%)
Overall	**All**	**−3.9**	**−1.5**	**−1.4**	**−2.7**	**−0.7**	**−0.6**	**−1.3**	**−0.6**	**−0.7**



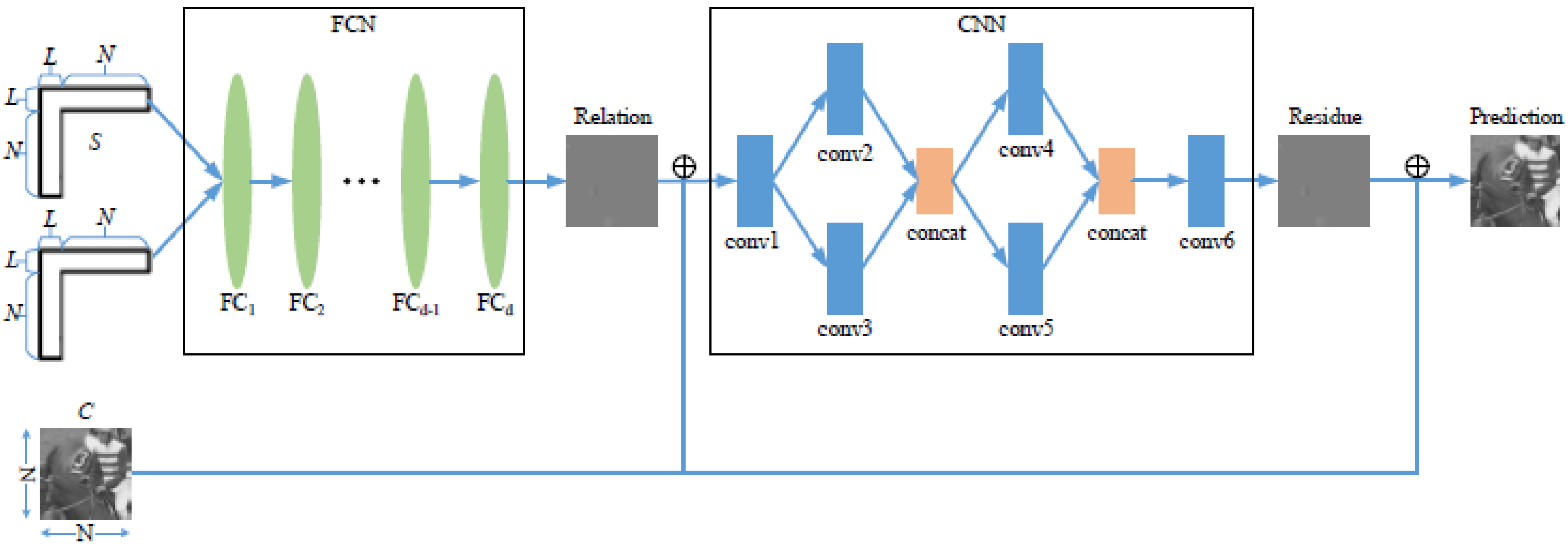
Class	Sequence	BD-rate of LDP			BD-rate of LDB			BD-rate of RA		
		Y	U	V	Y	U	V	Y	U	V
All Sequences	Overall	-2.2%	-0.6%	-0.5%	-1.2%	0.4%	-0.1%	-0.9%	-0.3%	-0.6%

Yan, N., Liu, D., Li, H., Li, B., Li, L., Wu, F. Convolutional Neural Network-Based Fractional-Pixel Motion Compensation. IEEE Transactions on Circuits and Systems for Video Technology, 2018.

Liu, J., Xia, S., Yang, W., Li, M., Liu, D. One-for-All: Grouped Variation Network-Based Fractional Interpolation in Video Coding. IEEE Transactions on Image Processing, 28(5), 2140-2151, 2019.

帧间预测增强 (1)

- 利用空域-时域联合信息
  - 空域周边重建像素
  - 时域参考像素

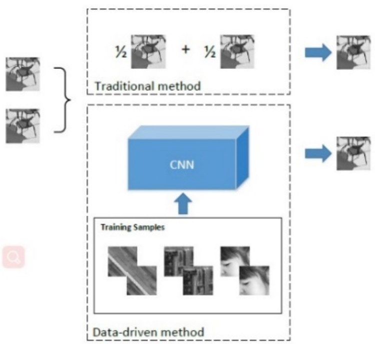


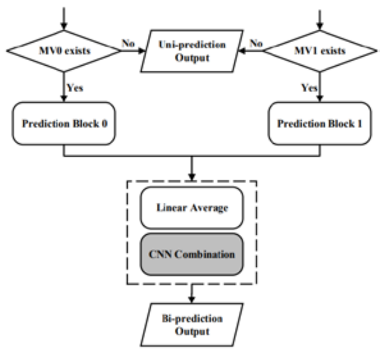
Y. Wang et al. Neural network based inter prediction for HEVC, IEEE ICME, 2018.

帧间预测增强(2)

提升预测准确性

帧间双向预测预测（BIP）



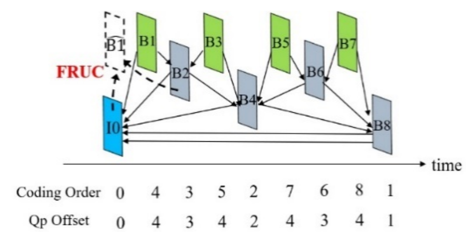


Sequences	BIP-CNN vs. HM-16.9	
	RA	LDB
Class A	-2.1 %	-1.7 %
Class B	-3.2 %	-1.9 %
Class C	-2.2 %	-0.9 %
Class D	-3.2 %	-1.0 %
Class E	/	-2.8 %
Overall	-2.7 %	-1.7 %
Encode Time	149%	185%
Decode Time	4259%	2853%

Zhao, Z., Wang, S., Wang, S., Zhang, X., Ma, S., Yang, J. Enhanced bi-prediction with convolutional neural network for high efficiency video coding. *IEEE Transactions on Circuits and Systems for Video Technology, 2018*.

虚拟参考帧生成

Deep Virtual Reference Frame, DVRF



Sequences	DVRF	
	RA (HM16.9)	RA (JEM7.1)
Class A	-6.7%	-1.3%
Class B	-3.5%	-0.4%
Class C	-4.0%	-0.8%
Class D	-5.7%	-0.7%
Class E	/	-0.8%
Overall	-4.6%	-0.7%
Encode Time	135%	124%
Decode Time	4376%	1025%

Zhao, L., Wang, S., Zhang, X., Wang, S., Ma, S., Gao, W. Enhanced CTU-level inter prediction with deep frame rate up-conversion for high efficiency video coding. *IEEE International Conference on Image Processing*, pp. 206-210, 2018.

环路滤波

基于整帧的处理：SRCNN

- In-loop

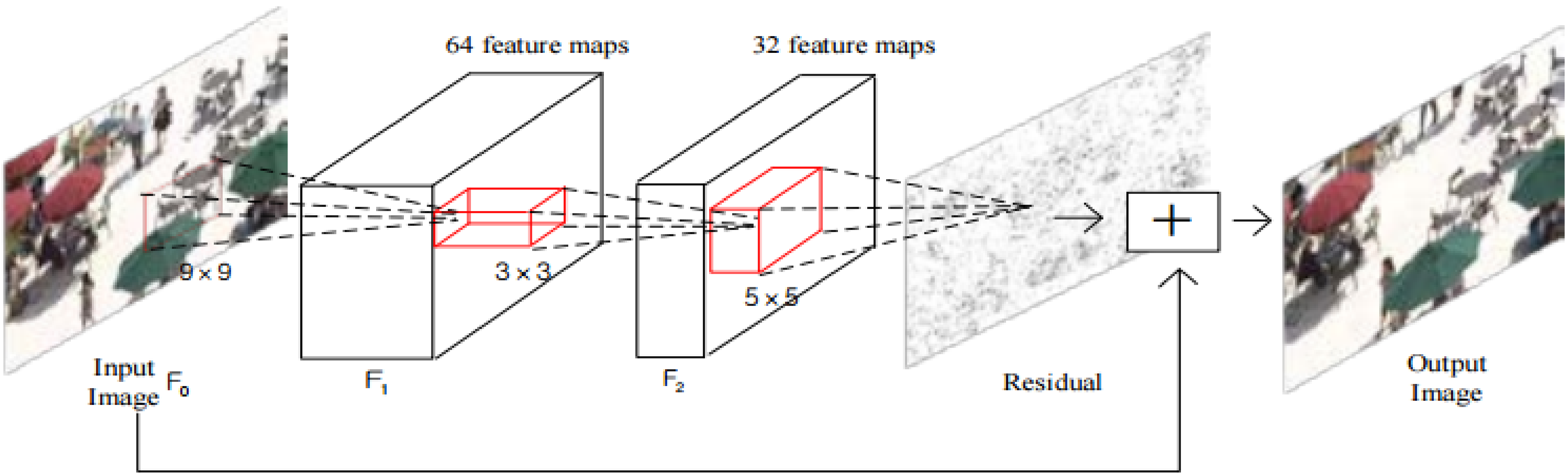


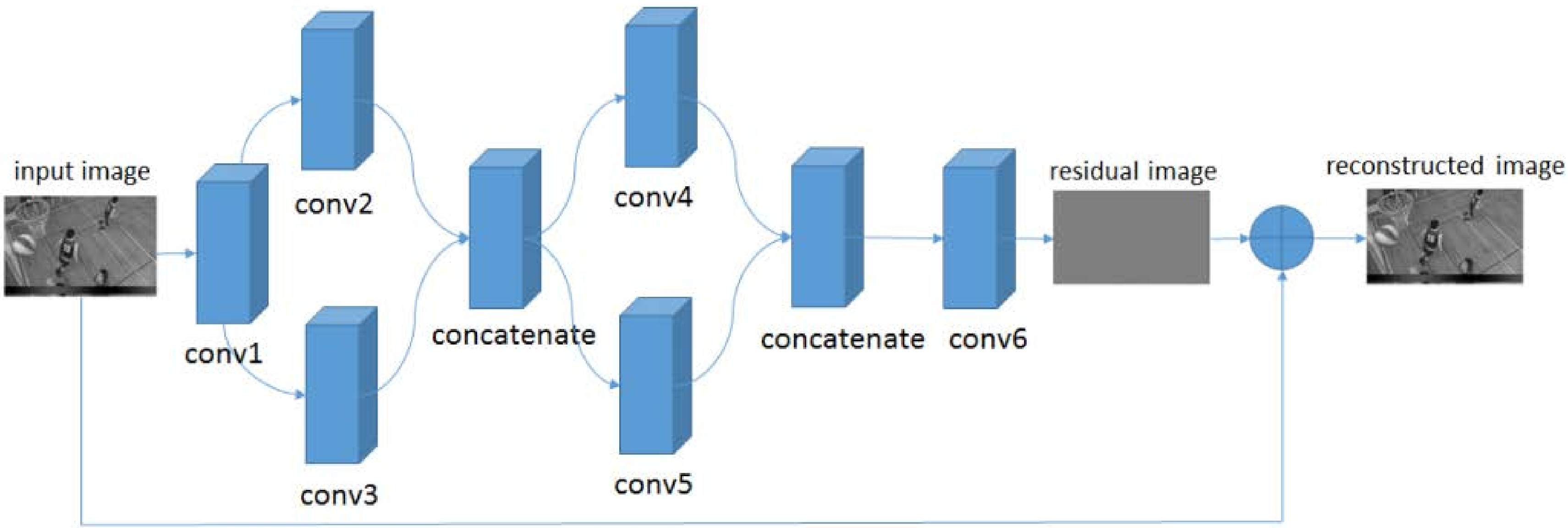
Table 1. Performance of our proposed IFCNN in comparison with SAO in terms of BD rates (BDBR).

Sizes	Seq.	All Intra	LDP-Case I	LDP-Case II	RA-Case I	RA-Case II
		BDBR (%)	BDBR (%)	BDBR (%)	BDBR (%)	BDBR (%)
832×480	BD	-10.1	-5.3	-3.0	-6.0	-6.7
	BQM	-3.7	-3.0	-2.4	-2.4	-2.9
	PS	-2.7	-2.0	-1.2	0.0	-1.1
	BDT	-7.6	-3.5	-2.4	-4.3	-4.9
416×240	BP	-3.3	-2.8	-1.5	-0.6	-1.1
	BQS	-2.4	-3.3	-2.9	1.4	-0.8
	B	-3.4	-2.3	-2.6	0.0	-1.4
	RH	-4.9	-0.4	0.6	-1.2	-1.6
	Avg.	-4.8	-2.8	-1.9	-1.6	-2.6

Park W S, Kim M. CNN-based in-loop filtering for coding efficiency improvement. IEEE Image, Video, and Multidimensional Signal Processing Workshop (IVMSP) 2016: 1-5.

环路滤波

- 帧内编码后处理
  - Post processing, All Intra
  - QP: 22, 27, 32, 37



Dai Y, Liu D, Wu F. A Convolutional Neural Network Approach for Post-Processing in HEVC Intra Coding. MMM 2017: 28-39.

环路滤波

Class	Sequence	BD-rate		
		Y (%)	U (%)	V (%)
Class A	Traffic	-5.6	-3.5	-4.1
	PeopleOnStreet	-5.4	-5.9	-5.7
	Nebuta	-0.9	-4.9	-4.1
	SteamLocomotive	-1.9	-0.5	-0.3
Class B	Kimono	-2.5	-1.5	-1.4
	ParkScene	-4.4	-3.3	-2.5
	Cactus	-4.6	-3.9	-6.3
	BasketballDrive	-2.5	-3.7	-5.3
	BQTerrace	-2.6	-3.3	-3.0
Class C	BasketballDrill	-6.9	-5.8	-6.8
	BQMall	-5.1	-5.3	-5.3
	PartyScene	-3.6	-4.4	-4.4
	RaceHorses	-4.2	-6.7	-11.0
Class D	BasketballPass	-5.3	-4.4	-6.5
	BQSquare	-3.8	-4.2	-6.4
	BlowingBubbles	-4.9	-8.4	-7.9
	RaceHorses	-7.6	-8.5	-11.5
Class E	FourPeople	-7.0	-5.3	-5.2
	Johnny	-5.9	-5.0	-5.5
	KristenAndSara	-6.7	-6.1	-6.2
Class Summary	Class A	-3.5	-3.7	-3.6
	Class B	-3.3	-3.2	-3.7
	Class C	-5.0	-5.5	-6.9
	Class D	-5.4	-6.4	-8.1
	Class E	-6.5	-5.5	-5.6
Overall	**All**	**-4.6**	**-4.7**	**-5.5**

Performance

Left: HEVC CTC,

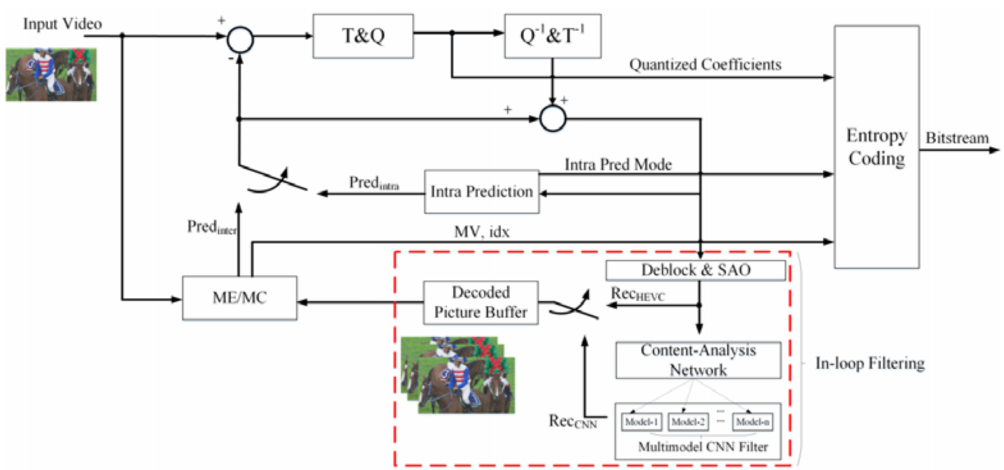
Right: Compare with other networks

Network		BD-rate		
		Y (%)	U (%)	V (%)
AR-CNN	Class A	0.9	2.1	2.1
	Class B	1.0	3.3	4.5
	Class C	-0.6	2.6	4.0
	Class D	-0.8	1.9	2.0
	Class E	0.4	5.5	6.1
	Overall	**0.2**	**3.0**	**3.7**
VDSR	Class A	-2.8	-3.2	-3.1
	Class B	-2.7	-2.7	-3.3
	Class C	-4.1	-4.8	-5.7
	Class D	-4.4	-5.6	-7.3
	Class E	-5.7	-5.7	-6.1
	Overall	**-3.8**	**-4.3**	**-4.9**

环路滤波

基于内容特性的神经网络环路滤波

- 针对不同内容特性的视频图像训练CNN模型
- Content analysis + CNN in-loop filter



设计CNN环路滤波语法元素

MMCNN_parameter_Set()	Descriptor
Y_MMCNN_on	1 bit
Cb_MMCNN_on	1 bit
Cr_MMCNN_on	1 bit
if(Y_MMCNN_on){	
for (i=0;i¡NumOfCTUs; ++i){	
CTU_MMCNN_on[i]	1 bit
}	
}	

Sequences	AI			LDB			LDP			RA		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
Class A	-4.7%	-3.3%	-2.6%	-6.7%	-2.6%	-1.9%	-3.5%	0.2%	0.3%	-6.6%	-3.4%	-3.0%
Class B	-3.5%	-2.8%	-3.0%	-5.7%	-1.6%	-2.2%	-4.5%	-0.5%	-1.1%	-6.5%	-2.5%	-2.7%
Class C	-3.4%	-3.5%	-5.0%	-5.0%	-3.4%	-5.0%	-4.4%	-1.9%	-3.0%	-4.5%	-3.3%	-4.5%
Class D	-3.2%	-4.7%	-6.0%	-3.8%	-1.7%	-2.6%	-3.5%	-0.8%	-0.9%	-3.3%	-2.6%	-3.6%
Class E	-5.8%	-4.1%	-5.2%	-8.6%	-5.2%	-5.6%	-7.7%	-1.7%	-0.9%	-9.0%	-4.2%	-5.3%
Overall	-4.1%	-3.7%	-4.1%	-6.0%	-2.9%	-3.5%	-4.7%	-1.0%	-1.2%	-6.0%	-3.2%	-3.8%

Jia, C., Wang, S., Zhang, X., Wang, S., Liu, J., Pu, S., Ma, S. Content-Aware Convolutional Neural Network for In-loop Filtering in High Efficiency Video Coding. *IEEE Transactions on Image Processing, 2019*.

环路滤波

- 智能编码与VVC（JVET-N0110）
  - ALF之后
  - 浅层卷积神经网络
  - 亮度分量与色度分量共享网络
  - 帧级开关、CTU级开关、32x32块开关

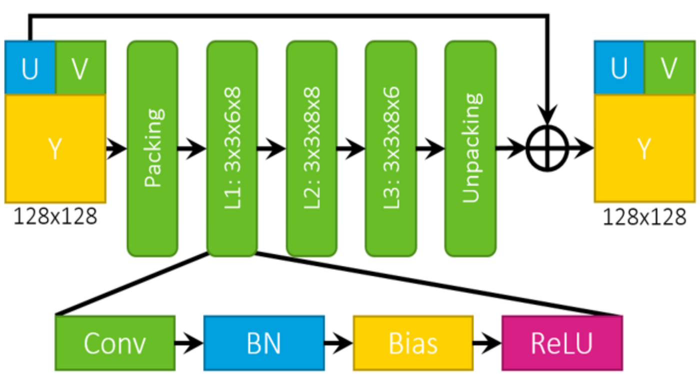
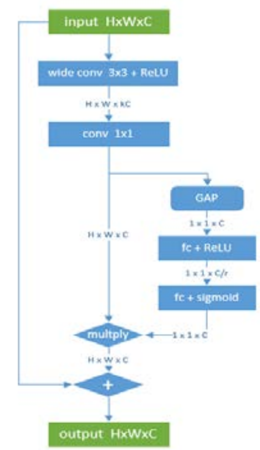


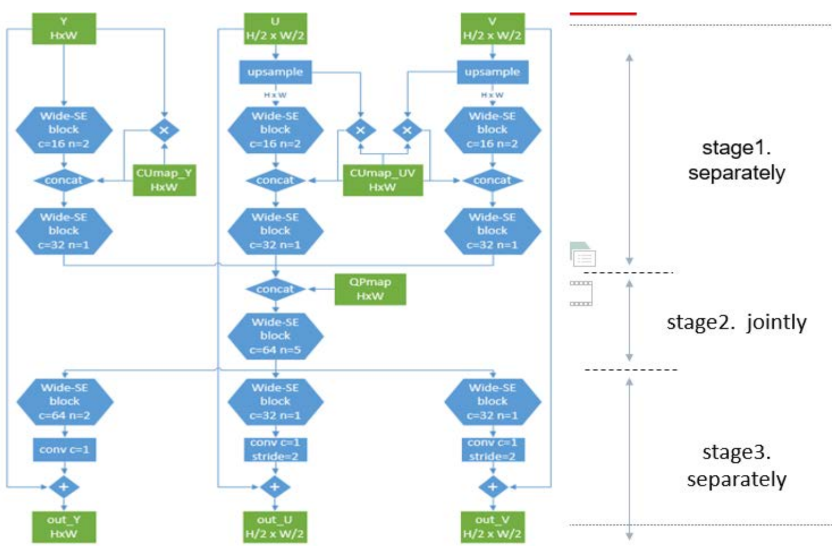
Table 3. Results of the proposed CNNLF

	Random access Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1	-2.95%	-8.57%	-13.33%	100%	189%
Class A2	-1.47%	-18.33%	-15.72%	100%	130%
Class B	-1.52%	-23.99%	-21.70%	100%	148%
Class C	0.12%	-5.94%	-6.99%	99%	116%
Class E					
Overall	-1.36%	-14.96%	-14.91%	100%	142%
Class D					
Class F					

环路滤波

- 智能编码与VVC（JVET-N0133）
  - 代替Deblock, SAO, ALF
  - 辅助输入信息：块划分结构和QP
  - SE(Squeeze and Excitation) block

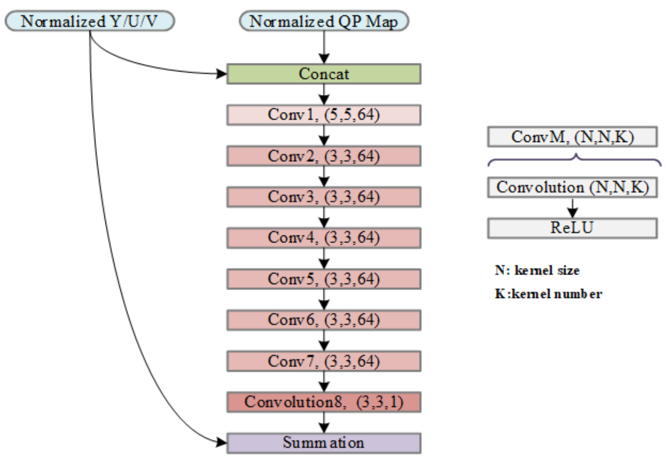




	All Intra Main10 (GPU)				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Class A1	#VALUE!	#VALUE!	#VALUE!	#NUM!	#NUM!
Class A2	#VALUE!	#VALUE!	#VALUE!	#NUM!	#NUM!
Class B	-3.99%	-10.41%	-10.75%	81%	1762%
Class C	-5.92%	-9.56%	-12.46%	61%	2847%
Class E	-7.26%	-6.89%	-9.10%	60%	2451%
Overall	-5.72%	-8.95%	-10.77%	67%	2353%
Class D	-6.24%	-12.31%	-16.61%	62%	4510%
Class F	10.52%	11.89%	6.30%	62%	2127%

环路滤波

- 智能编码与VVC（JVET-N0169）
  - CNNLF的位置
  - 辅助输入信息：QP Map
  - 并行化：分块滤波



Results for CNNF before the SAO

	All Intra Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Overall	-3.48%	-5.18%	-6.77%	142%	38414%

Results for CNNF before the ALF with DF and SAO turned off

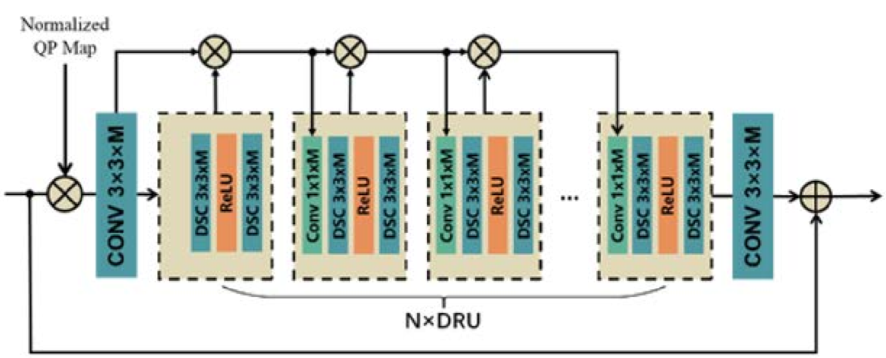
	All Intra Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Overall	-4.65%	-6.73%	-7.92%	139%	37956%

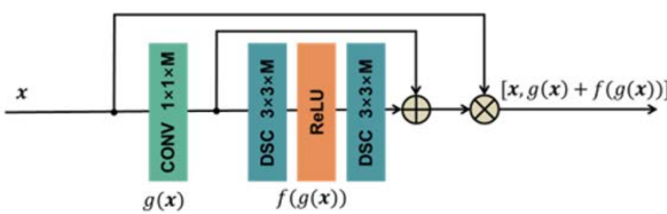
Results for CNNF without all the conventional filters

	All Intra Main10				
	Over VTM-4.0				
	Y	U	V	EncT	DecT
Overall	-4.14%	-5.49%	-6.70%	140%	38411%

环路滤波

- 智能编码与VVC（JVET-N0254）
  - Dense Residual CNN
  - 深度可分离卷积(DSC)减少参数量





All Intra Main10 Over VTM-4.0	Y	U	V	EncT	DecT
Class A1	-0.97%	-1.64%	-2.91%	114%	5904%
Class A2	-1.46%	-2.72%	-1.77%	107%	3465%
Class B	-0.93%	-2.19%	-3.08%	106%	3665%
Class C	-1.90%	-2.34%	-3.33%	104%	4520%
Class E	-2.57%	-1.57%	-2.13%	108%	7759%
Overall	-1.52%	-2.12%	-2.73%	107%	4667%
Class D	-2.22%	-0.92%	-3.37%	103%	3955%

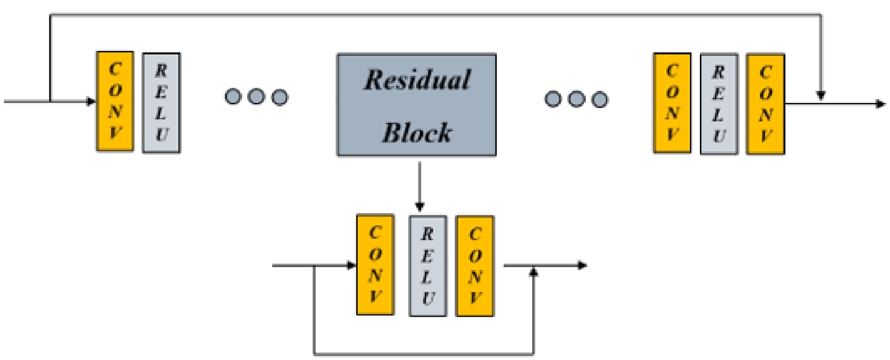
Random Access Main10 Over VTM-4.0	Y	U	V	EncT	DecT
Class A1	-1.27%	-3.38%	-5.10%	106%	6967%
Class A2	-2.21%	-5.74%	-2.88%	106%	6435%
Class B	-1.13%	-4.73%	-4.55%	106%	7011%
Class C	-1.39%	-3.63%	-4.36%	106%	8110%
Class E					
Overall	-1.45%	-4.37%	-4.27%	106%	7156%
Class D	-1.39%	-1.96%	-3.08%	105%	4217%

Low delay B Main10 Over VTM-4.0	Y	U	V	EncT	DecT
Class A1					
Class A2					
Class B	-1.12%	-7.24%	-7.12%	105%	8387%
Class C	-1.40%	-6.97%	-6.45%	104%	9894%
Class E	-2.45%	-2.80%	-2.95%	116%	9434%
Overall	-1.54%	-6.04%	-5.86%	108%	9127%
Class D	-1.73%	-2.75%	-4.92%	105%	6881%

环路滤波

- 智能编码与AVS3
  - QP分段训练残差网络
  - 代替Deblock, SAO, ALF
  - 帧级开关、CTU级开关

序列（4K）	CNN Loop Filter vs HPM-2.4, AI (4K)		
	Y	U	V
Campfire_3840x2160_30	-5.73%	-10.29%	-2.45%
DaylightRoad2_3840x2160_60	-7.60%	-0.67%	-4.42%
ParkRunning3	-3.81%	-3.73%	-3.42%
Tango2_3840x2160_60	-4.52%	-3.41%	-5.14%
平均性能	**-5.41%**	**-4.53%**	**-3.85%**



AVS-M4730

序列（1080p 720p）	CNN Loop Filter vs HPM-2.4, AI			CNN Loop Filter vs HPM-2.4, RA		
	Y	U	V	Y	U	V
BasketballDrive_1920x1080_50	-8.03%	-11.57%	-15.57%	-2.51%	-2.59%	-0.11%
Cactus_1920x1080_50	-7.69%	-6.37%	-8.04%	-6.43%	-6.70%	-7.93%
City_1280x720_60	-3.61%	-4.28%	-3.17%	-2.53%	-5.55%	-6.00%
Crew_1280x720_60	-4.12%	-2.67%	-4.06%	-3.12%	-6.50%	-9.10%
Vidyo1_1280x720_60	-11.29%	-4.65%	-8.56%	-8.87%	-5.88%	-8.63%
Vidyo3_1280x720_60	-8.45%	-3.19%	-0.97%	-1.53%	-2.94%	-6.15%
平均性能	**-7.20%**	**-5.45%**	**-6.74%**	**-4.17%**	**-5.03%**	**-6.32%**

帧内编码模式决策

- 基于CNN的CU模式决策
  - 1. 分析CU块纹理
  - 2. 减少CU模式的数目
  - 3. 引入QP作为辅助信息

Liu Z, Yu X, Chen S, et al. CNN oriented fast HEVC intra CU mode decision. ISCAS 2016: 2270-2273.

帧内编码模式决策

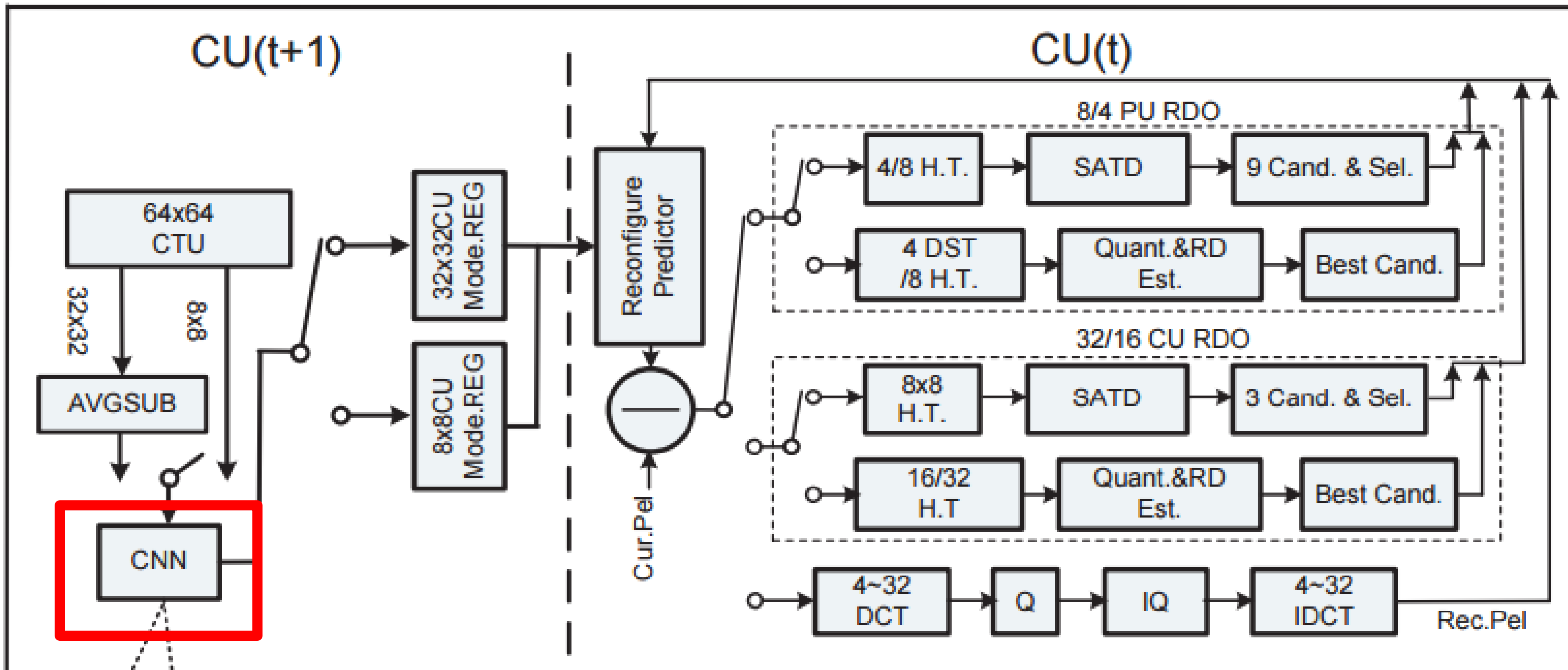
- 实现
  - Adding FastCUMode()
  - into xCompressCU

function XCompressCU(*pCurCU)
    $m \leftarrow$ FastCUMode(pCurCU,QP,QS)
    if $m \neq$SPLIT **then**
        $C_{2N} \leftarrow$CheckINTRA(pCurCU)
    else
        $C_{2N} \leftarrow \infty$
    end if
    if $m \neq$ HOMO and $curD < maxD$ **then**
        $C_N \leftarrow 0$
        for $i = 0$ to 3 **do**
            pSubCU$_i \leftarrow$ pointer to SubCU$_i$
            $C_N \leftarrow C_N +$ XCompressCU(pSubCU$_i$)
        end for
    else
        $C_N \leftarrow \infty$
    end if
    CheckBestMode(C_{2N}, C_N)
end function

Liu Z, Yu X, Chen S, et al. CNN oriented fast HEVC intra CU mode decision. ISCAS 2016: 2270-2273.

帧内编码模式决策

- HEVC Intra 硬件编码器实现
  - Big / Small CU pipeline



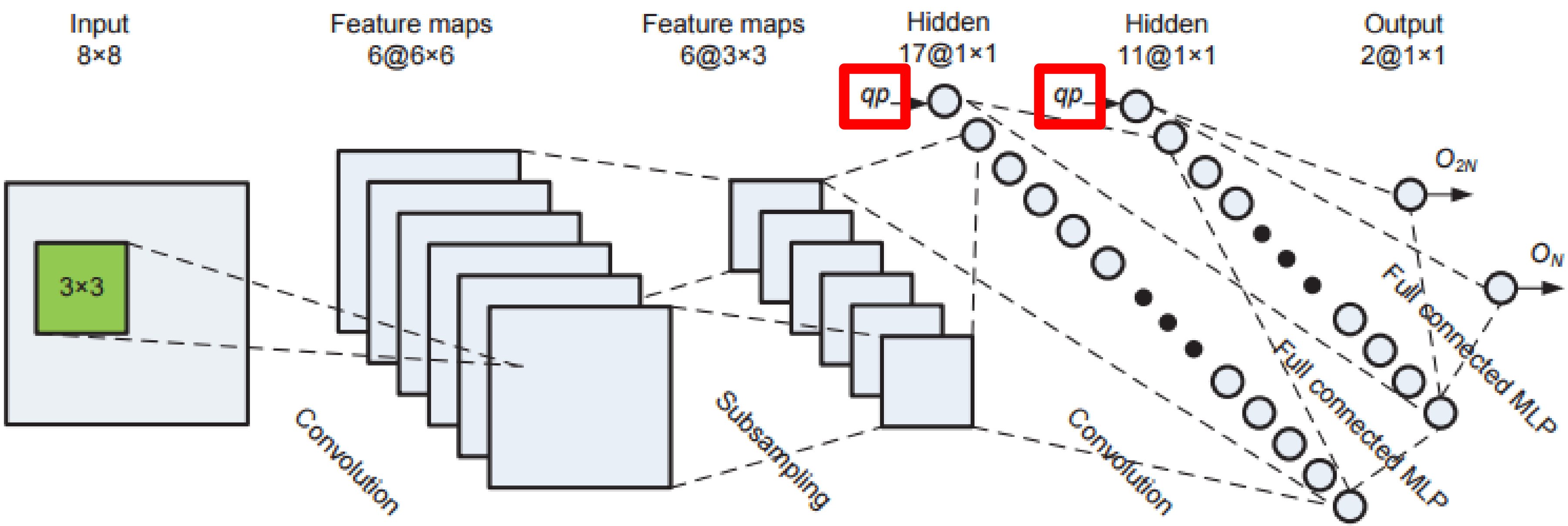
(a). Top Block Diagram of Proposed HEVC Intra Encoder

Liu Z, Yu X, Chen S, et al. CNN oriented fast HEVC intra CU mode decision. ISCAS 2016: 2270-2273.

帧内编码模式决策

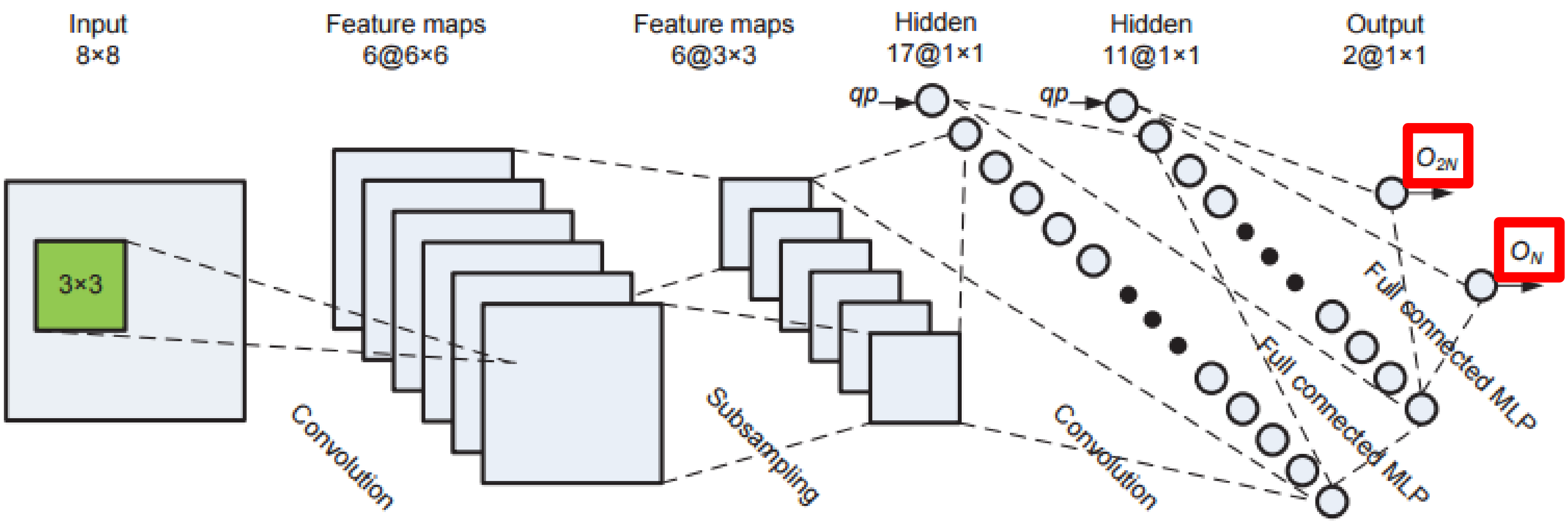
- 使用类似LeNet结构

Taking QP into consideration



(b) Architecture of Convolutional Neural Network for CU Mode Decision

帧内编码模式决策

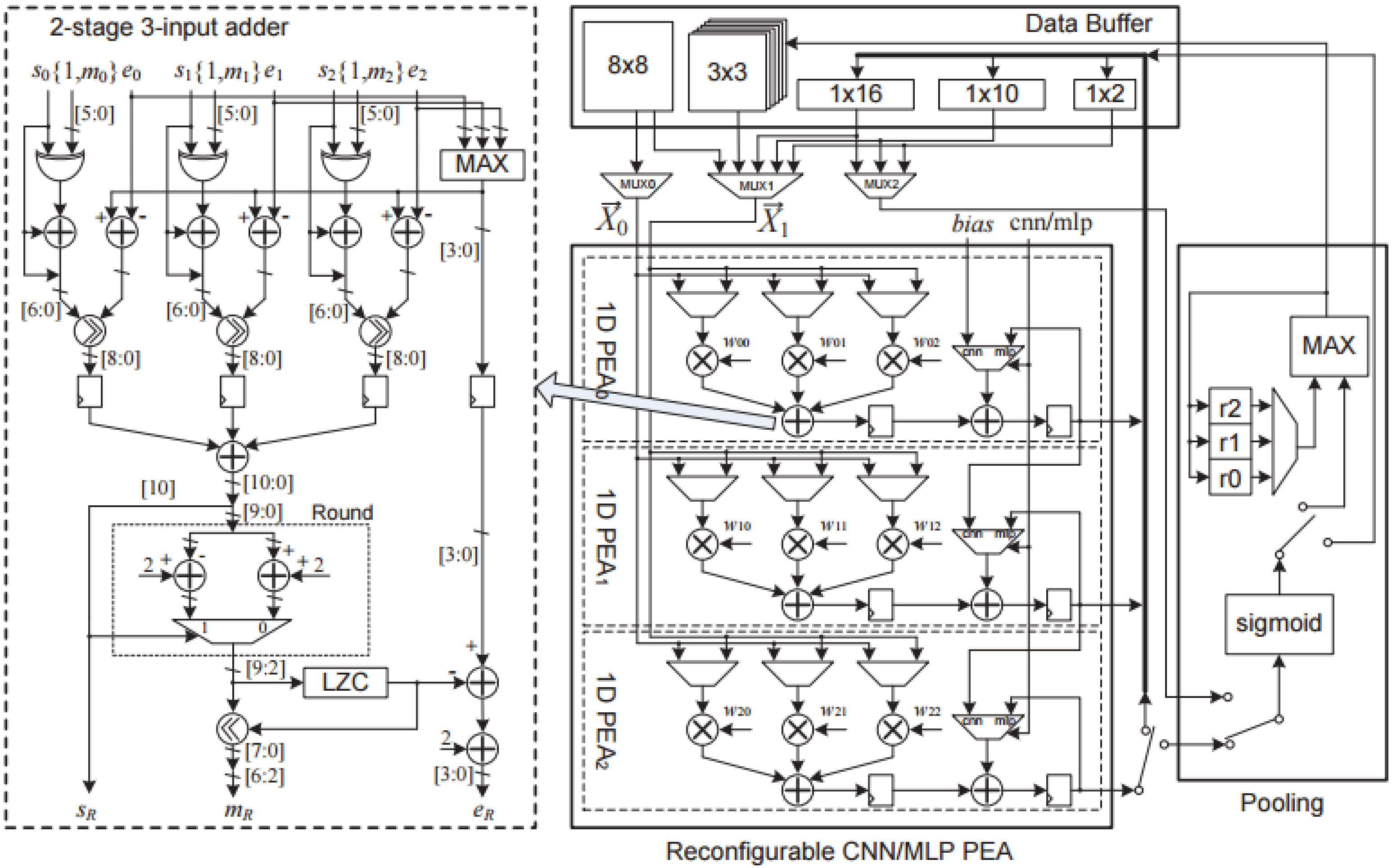


(b) Architecture of Convolutional Neural Network for CU Mode Decision

- 将编码模式决策建模为二分类问题
  - 预测当前编码单元是否划分

帧内编码模式决策

- VLSI 设计CNN加速模块



帧内编码模式决策

- 性能对比
  - 63% time save with 2.7% loss in BDBR

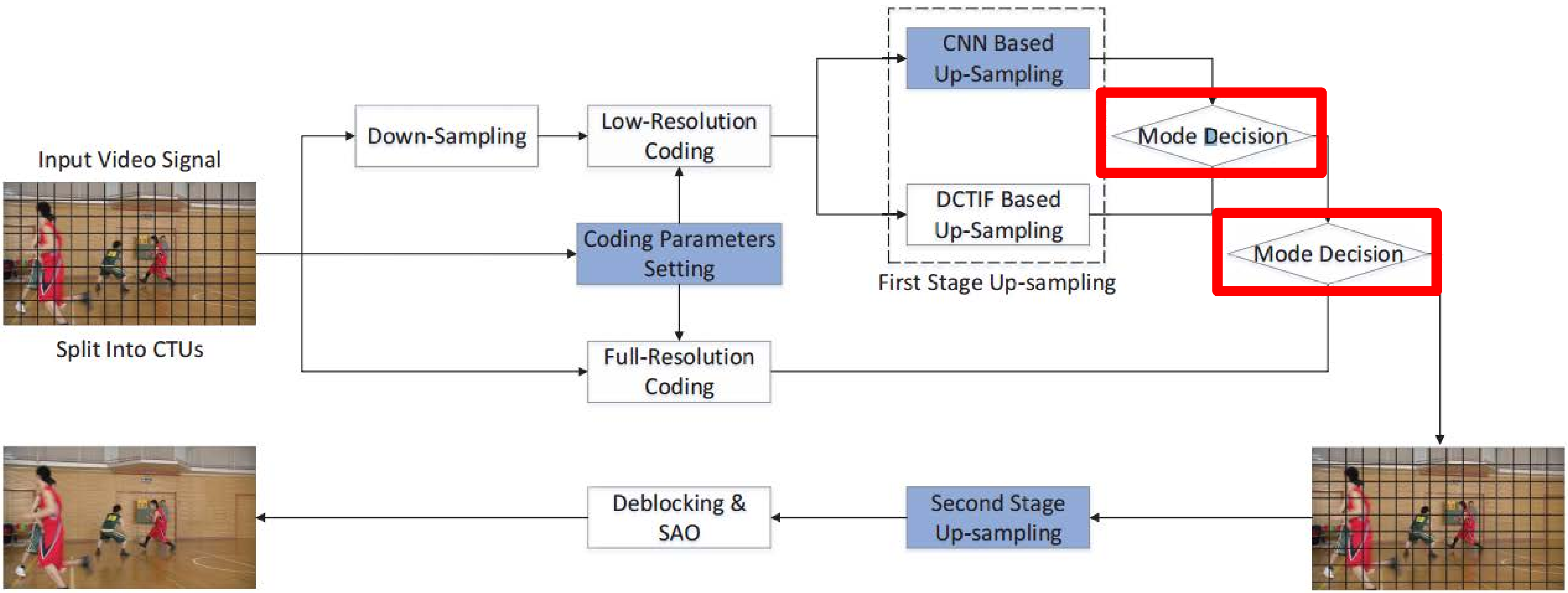
Table III: Performance Comparison between Proposed Solution and Existing Algorithms

Algorithm	$\Delta T_{CMD}[\%]$	$\Delta T_{PMD}[\%]$	BDBR[%]	$\Delta T[\%]$	VLSI
[3]	50-α	α	0.7	50	No
[5]†	26	45	1.0	60	No
[6]	52	0	0.8	52	No
[7]†	52	5	5.1	57	Yes
[8]	62	0	4.5	62	Yes
Proposed	63	0	2.7	63	Yes

† indicates that class F sequences were not tested.

下采样-上采样编码

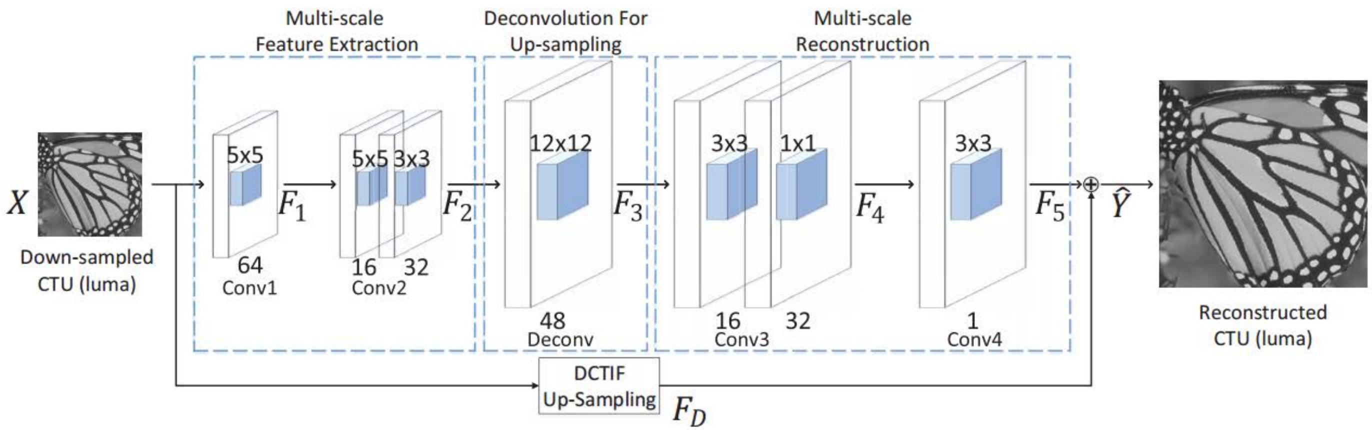
- CTU级处理
  - 两级RDO
    - 1. 是否采用变分辨率编码
    - 2. 上采样模块使用DCT插值或CNN



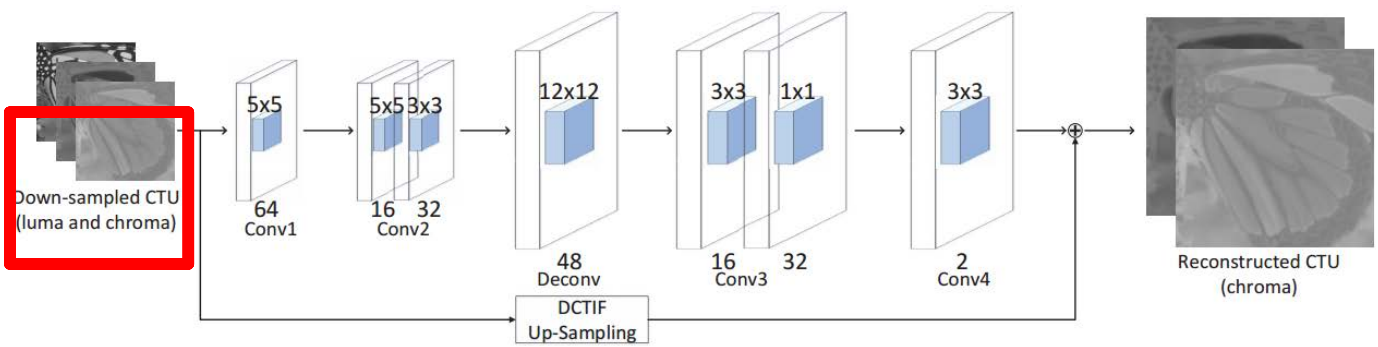
Li Y, Liu D, Li H, Li L, Wu F. Convolutional Neural Network-Based Block Up-sampling for Intra Frame Coding.

下采样-上采样编码

- 亮度分量网络
  - Input: low resolution patch, output: high resolution



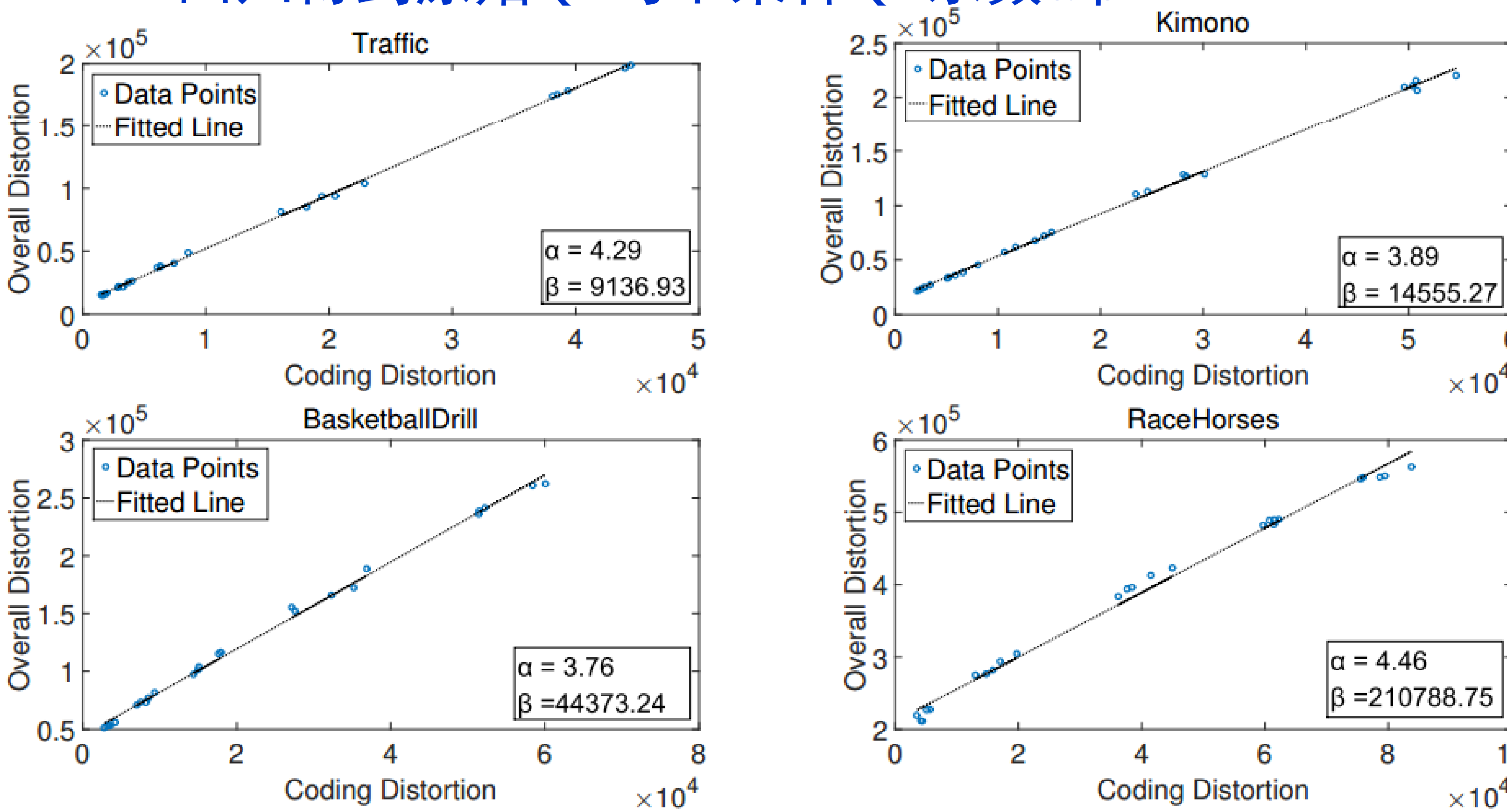
- 色度分量处理：使用亮度作为引导



Li Y, Liu D, Li H, Li L, Wu F. Convolutional Neural Network-Based Block Up-sampling for Intra Frame Coding.

下采样-上采样编码

- 下采样后应调整QP
  - 回归得到原始QP与下采样QP系数α β



Li Y, Liu D, Li H, Li L, Wu F. Convolutional Neural Network-Based Block Up-sampling for Intra Frame Coding.

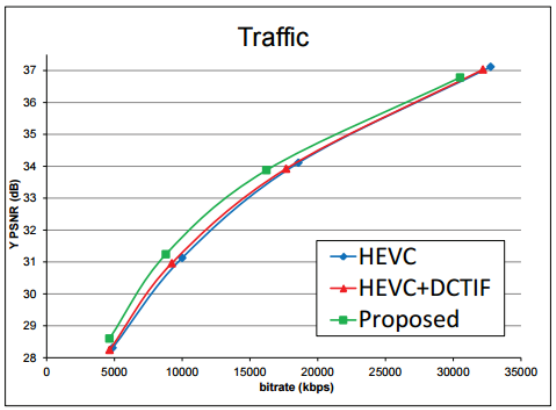
下采样-上采样编码

- 测试条件
  - Qp: 32, 37, 42, 47

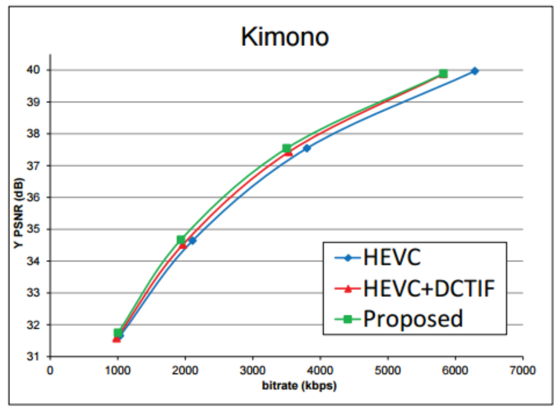
Class	Sequence	BD-Rate (Anchored on HEVC)				BD-Rate (Anchored on HEVC+DCTIF)			
		Y	U	V	Y SSIM	Y	U	V	Y SSIM
Class A	Traffic	–10.1%	–3.5%	6.0%	–12.9%	–8.0%	–13.2%	–2.6%	–7.9%
	PeopleOnStreet	–9.7%	–14.8%	–14.5%	–12.9%	–8.5%	–20.4%	–18.5%	–9.7%
	Nebuta	–2.0%	–22.0%	3.1%	–4.4%	–1.7%	–22.5%	1.6%	–3.6%
	SteamLocomotive	–1.7%	–27.7%	–25.4%	–6.1%	–1.2%	–34.2%	-25.6%	–2.8%
Class B	Kimono	–7.7%	–5.5%	18.8%	–9.6%	–3.4%	–25.9%	–4.3%	–3.4%
	ParkScene	–7.1%	–14.4%	–2.3%	–11.3%	–5.0%	–25.2%	–14.6%	–6.6%
	Cactus	–6.6%	–2.5%	8.3%	–10.0%	–5.0%	–6.5%	0.9%	–6.7%
	BQTerrace	–3.7%	–7.6%	–9.1%	–9.6%	–3.1%	–8.2%	–7.1%	–6.5%
	BasketballDrive	–6.1%	–1.2%	3.2%	–10.8%	–3.4%	–5.8%	–2.5%	–3.8%
Class C	BasketballDrill	–4.9%	4.5%	8.1%	–7.9%	–4.0%	4.9%	2.1%	–6.6%
	BQMall	–2.9%	–7.2%	–7.2%	–6.2%	–2.3%	–10.6%	–9.1%	–5.3%
	PartyScene	–1.0%	–5.1%	–1.6%	–4.0%	–1.0%	–5.5%	–3.2%	–3.6%
	RaceHorsesC	–6.7%	4.6%	7.5%	–10.7%	–6.0%	1.9%	3.9%	–8.6%
Class D	BasketballPass	–2.0%	–3.7%	9.2%	–4.3%	–2.3%	–7.5%	12.3%	–4.4%
	BQSquare	–0.9%	–0.6%	–21.1%	–1.4%	–0.5%	1.7%	-16.7%	–1.2%
	BlowingBubbles	–3.2%	3.1%	–8.0%	–5.3%	–1.7%	0.5%	-9.6%	–3.8%
	RaceHorses	–9.9%	7.5%	6.4%	–12.6%	–9.6%	5.0%	6.6%	–11.1%
Class E	FourPeople	–7.2%	–10.5%	–11.0%	–11.0%	–7.2%	–14.7%	–14.5%	–9.5%
	Johnny	–9.0%	–3.2%	–3.2%	–11.1%	–7.1%	–6.0%	–8.3%	–5.6%
	KristenAndSara	–6.8%	–11.2%	–11.1%	–13.0%	–5.3%	–8.4%	–10.6%	–8.2%
Class UHD	Fountains	–4.0%	–12.9%	–11.2%	–7.4%	–2.0%	–16.1%	–9.2%	–2.0%
	Runners	–11.2%	22.8%	–0.1%	–12.4%	–7.0%	0.9%	–13.7%	–6.0%
	Rushhour	–8.5%	4.4%	1.8%	–10.3%	–3.2%	–9.2%	–9.5%	–3.0%
	TrafficFlow	–12.7%	–11.7%	–5.8%	–12.7%	–6.9%	–17.3%	–11.9%	–5.6%
	CampfireParty	–8.4%	–10.8%	–0.8%	–9.5%	–6.5%	–10.8%	–5.0%	–6.4%
Average of Classes A-E		–5.5%	–6.0%	–2.2%	–8.8%	–4.3%	–10.0%	–6.0%	–5.9%
Average of Class UHD		–9.0%	–1.6%	–3.2%	–10.5%	–5.1%	–10.5%	–9.9%	–4.6%

下采样-上采样编码

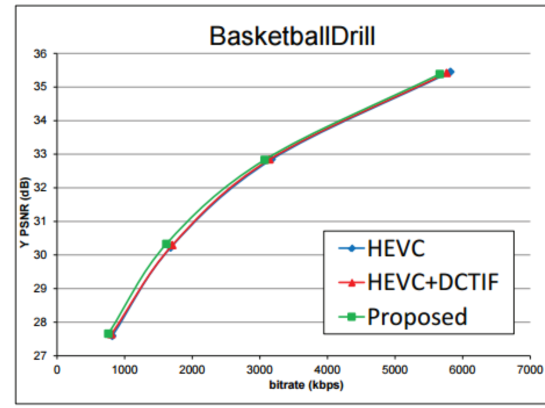
□ RD-曲线



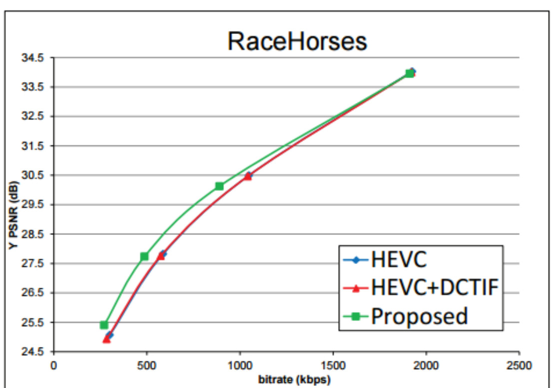
(a)



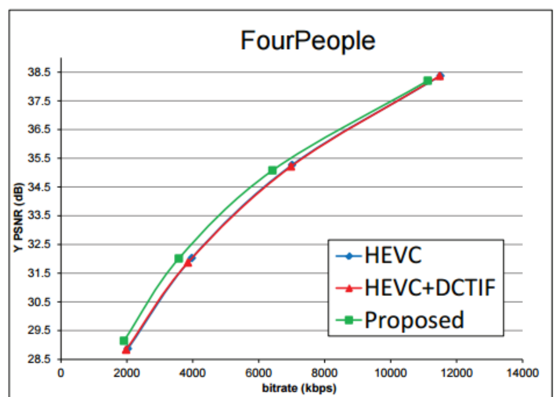
(b)



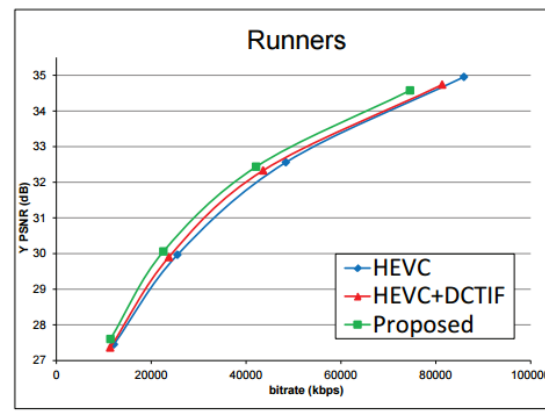
(c)



(d)



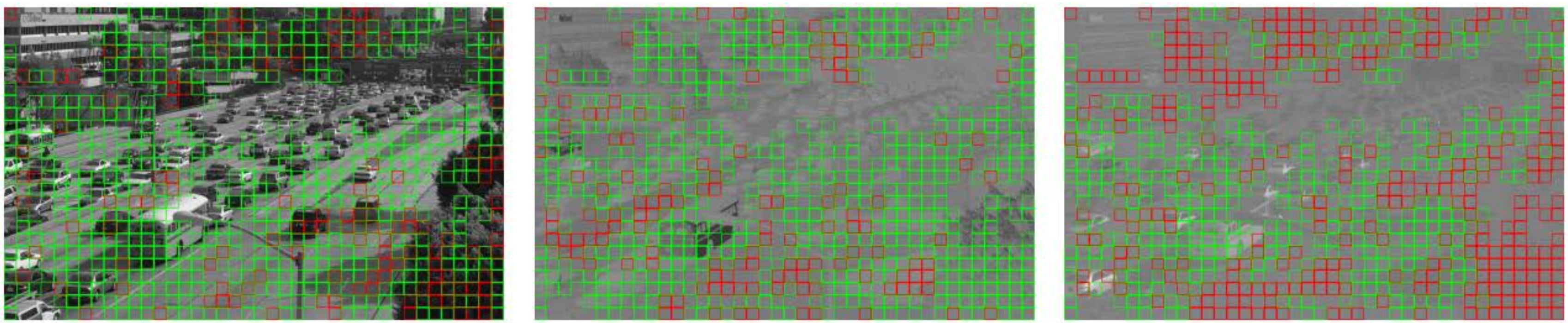
(e)



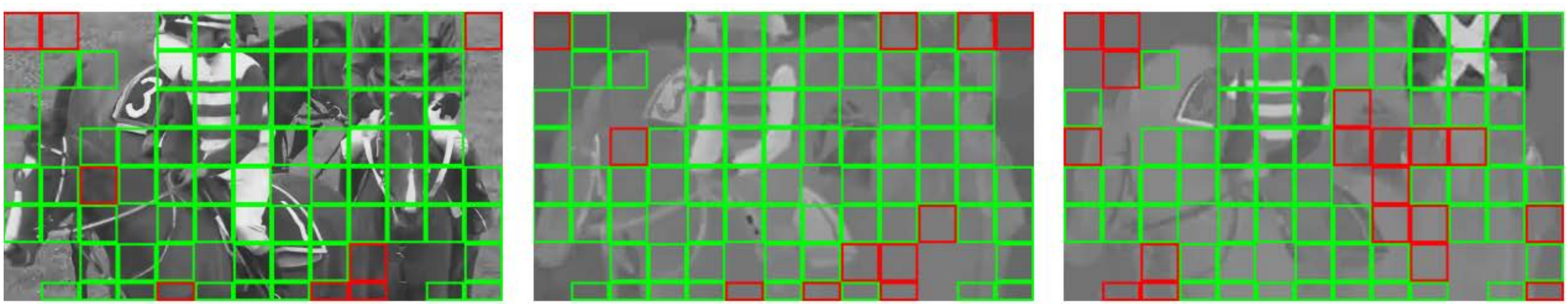
(f)

下采样-上采样编码

- 算法命中率
  - 绿色: 下采样编码 + CNN, 红色: 上采样编码 + DCTIF



QP 32



QP 42

目录

- 概述
- 神经网络视频编码历史
- 深度学习视频编码进展
- 展望

Future Works

- 深度学习为视频编码性能提升提供了新的思路
  - 未来可期
  - 有“大货”
  - 为什么能带来性能的明显提升值得进行理论探索
- 深度学习进入视频编码标准尚需进一步探索

参考文献

- [1] Cottrell G W, Munroe P, Zipser D. Image compression by neural network. An example of extensional programming[J]. Technical report, 1987.
- [2] Chua L O, Lin T. A neural network approach to transform image coding[J]. International journal of circuit theory and applications, 1988, 16(3): 317-324.
- [3] Luttrell S P. Image compression using a multilayer neural network[J]. Pattern Recognition Letters, 1989, 10(1): 1-7.
- [4] Luttrell S P. Image compression using a neural network[C]//Proc. IGARSS. 1988, 88: 1231-1238.
- [5] Naillon M, Theeten J B. Neural approach for TV image compression using a Hopfield type network[C]//Advances in neural information processing systems. 1989: 264-271.
- [6] Luttrell S P. The use of Bayesian and entropic methods in neural network theory[M]//Maximum entropy and Bayesian methods. Springer, Dordrecht, 1989: 363-370.
- [7] Sicuranza G L, Romponi G, Marsi S. Artificial neural network for image compression[J]. Electronics letters, 1990, 26(7): 477-479.
- [8] Kung S Y, Diamantaras K I. A neural network learning algorithm for adaptive principal component extraction (APEX)[C]//Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990 International Conference on. IEEE, 1990: 861-864.
- [9] Antonini M, Barlaud M, Mathieu P, et al. Multiscale image coding using the Kohonen neural network[C]//Visual Communications and Image Processing'90: Fifth in a Series. International Society for Optics and Photonics, 1990, 1360: 14-27.
- [10] Manikopoulos C N. Neural network approach to DPCM system design for image coding[J]. IEE Proceedings I (Communications, Speech and Vision), 1992, 139(5): 501-507.

参考文献

- [11] Jiang J. Image compression with neural networks–a survey[J]. Signal processing: image Communication, 1999, 14(9): 737-760.
- [12] Panchanathan S, Yeap T H, Pilache B. Neural network for image compression[C]//Applications of Artificial Neural Networks III. International Society for Optics and Photonics, 1992, 1709: 376-386.
- [13] Jiang W W, Kiang S Z, Hakim N Z, et al. Lossless compression for medical imaging systems using linear/nonlinear prediction and arithmetic coding[C]//Circuits and Systems, 1993., ISCAS'93, 1993 IEEE International Symposium on. IEEE, 1993: 283-286.
- [14] Dony, Robert D., and Simon Haykin. "Neural network approaches to image compression." Proceedings of the IEEE 83.2 (1995): 288-303.
- [15] Doulamis N, Doulamis A, Kalogeras D, et al. Low bit-rate coding of image sequences using adaptive regions of interest[J]. IEEE transactions on circuits and systems for video technology, 1998, 8(8): 928-934.
- [16] Egmont-Petersen M, de Ridder D, Handels H. Image processing with neural networks—a review[J]. Pattern recognition, 2002, 35(10): 2279-2301.
- [17] Durai S A, Saro E A. Image compression with back-propagation neural network using cumulative distribution function[J]. World Academy of Science, Engineering and Technology, 2006, 17: 60-64.
- [18] Dony, Robert D., and Simon Haykin. "Neural network approaches to image compression." Proceedings of the IEEE 83.2 (1995): 288-303.
- [19] Namphol A, Chin S H, Arozullah M. Image compression with a hierarchical neural network[J]. IEEE transactions on Aerospace and Electronic Systems, 1996, 32(1): 326-338.
- [20] Jiang J. Image compression with neural networks–a survey[J]. Signal processing: image Communication, 1999, 14(9): 737-760.

参考文献

- [21] Venetianer, Peter L., et al. "Analogic CNN algorithms for some image compression and restoration tasks." IEEE transactions on circuits and systems I: Fundamental theory and applications 42.5 (1995): 278-284.
- [22] Ng K S, Cheng L M. Artificial neural network for discrete cosine transform and image compression[C]//Document Analysis and Recognition, 1997., Proceedings of the Fourth International Conference on. IEEE, 1997, 2: 675-678.
- [23] Venetianter P L, Roska T. Image compression by cellular neural networks[J]. IEEE transactions on circuits and systems I: fundamental theory and applications, 1998, 45(3): 205-215.
- [24] Papamarkos, Nikos, Charalambos Strouthopoulos, and Ioannis Andreadis. "Multithresholding of color and gray-level images through a neural network technique." Image and Vision Computing 18.3 (2000): 213-222.
- [25] Bakırcıoğlu H, Koçak T. Survey of random neural network applications[J]. European journal of operational research, 2000, 126(2): 319-330.
- [26] Clausen, Clifford, and Harry Wechsler. "Color image compression using PCA and backpropagation learning." pattern recognition 33.9 (2000): 1555-1560.
- [27] Watanabe, Eiji, and Katsumi Mori. "Lossy image compression using a modular structured neural network." Neural Networks for Signal Processing XI, 2001. Proceedings of the 2001 IEEE Signal Processing Society Workshop. IEEE, 2001.
- [28] Vilovic, Ivan. "An experience in image compression using neural networks." Multimedia Signal Processing and Communications, 48th International Symposium ELMAR-2006 focused on. IEEE, 2006.
- [29] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. science, 2006, 313(5786): 504-507.

谢谢！

Q & A