

JVET-U0094

EE: Tests on Neural Network-based In-Loop Filter

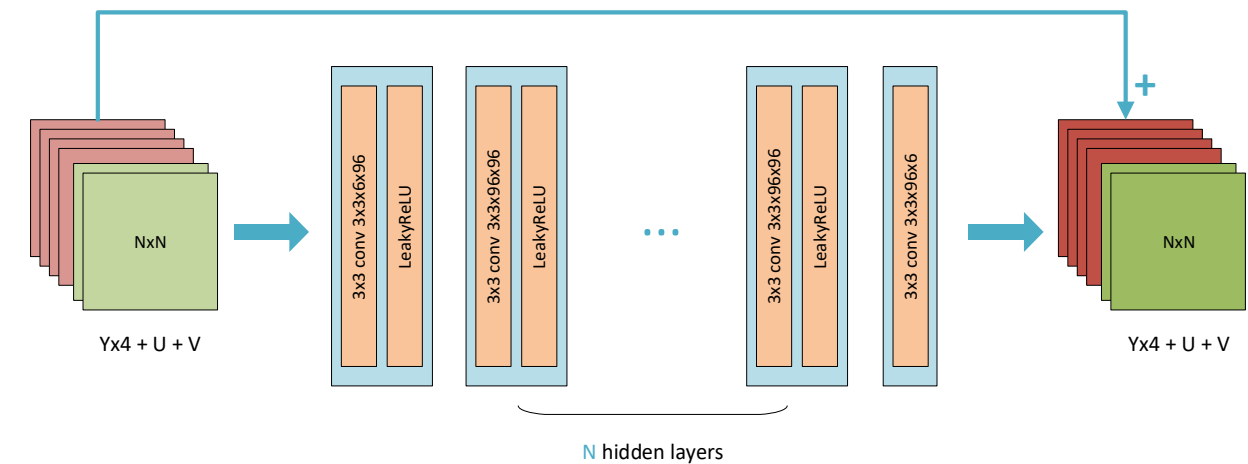
H. Wang, M. Karczewicz, J. Chen, A. M. Kotra



Network Structure

Proposed in JVET-T0079

- Input: a 128x128 YUV420 block
 - The input block is converted into a 64x64x6 tensor
 - Subtests 1~3 use QP info as an additional input plane
 - Pixel values are normalized to [0.0,1.0]
- Output: Residues to be added to input samples
- Number of conv. layers: 14 / 8
- Number of feature maps: 96
- Number of parameters: ~1M/~0.5M



Network Information in Inference Stage			
		N=12	N=6
Mandatory	HW environment:	CPU Only	
	GPU Type	Tensorflow v2.2.0	
	Framework:	0	
	Number of GPUs per Task		
	Total Parameter Number	1,007,814*4Models	509,574*4Models
	Parameter Precision (Bits)	32(F)	
	Memory Parameter (MB)	3.84 * 4 models	1.94 * 4 models
	MAC (Giga)	1.57e-4 Giga/pixel	7.93e-5 Giga/pixel
Optional	Total Conv. Layers	14	8
	Total FC Layers	0	0
	Batch size:	1	
	Patch size	64x64x6/64x64x7	
	Peak Memory Usage	1.5MB (input size 128x128)	

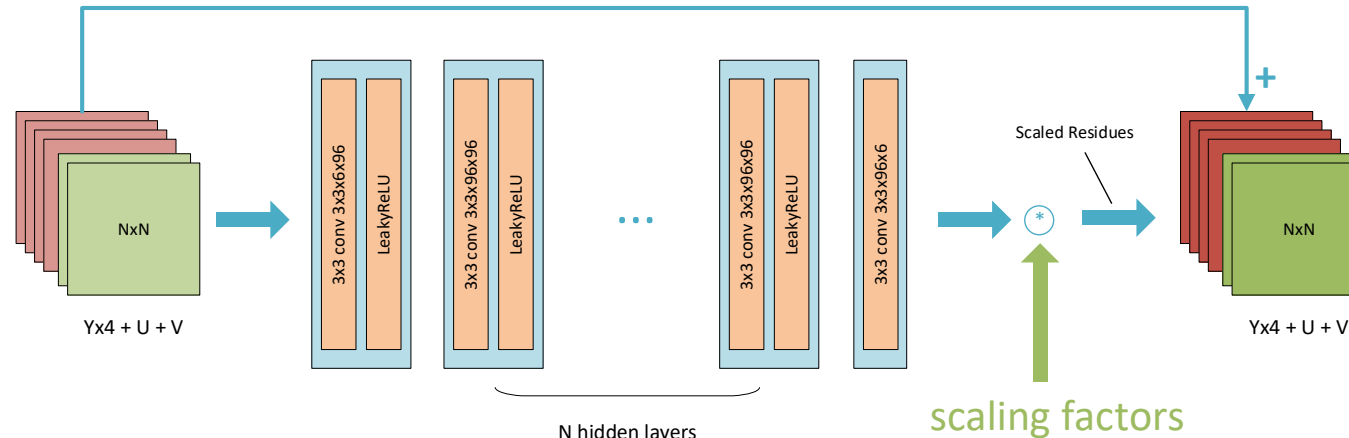
Training process

- Training data set: BVI-DVC
- Training data generation
 - QPs: 20, 25, 30, 35, 40, 45
 - Configurations: AI, RA
- Number of trained models:
 - 4 (2AI + 2RA) / 12 (6AI+6RA)

Network Information in Training Stage			
		N=12	N=6
Mandatory	GPU Type	GPU: NVIDIA Tesla V100-SXM2-32GB	
	Framework:	Tensorflow v2.2.0	
	Number of GPUs per Task	1	
	Epoch:	250	400
	Batch size:	10Kx64	10Kx64
	Training time:	104h	100h
	Training data information:	BVI-DVC	
	Training configurations for generating compressed training data (if different to VTM CTC):	QP=20,25,30,35,40,45 AI, RA	
Optional			
	Number of iterations	2.5M	4.0M
	Patch Size	64x64x6/64x64x7	
	Learning rate:	8.00E-04	
	Optimizer:	ADAM	
	Loss function:	Weighted L2	
	Preprocessing:	Convert 128x128 YUV420 signal to 6 64x64 blocks. Normalize to 0~1	
	Other information:		

VTM-10.0 integration

- Integrated as an additional filter prior SAO/after ALF
- Model selection is at picture level
- Slice/CTU level on/off is supported
- A scaling factor is signaled for each color component in the picture header



Test summary

- Aspects being tested in this EE:
 - Number of hidden layers (N=12, 6 are tested).
 - Number of candidate models for RD selection (X=4, 12 are tested).
 - Location of the NN filter (prior SAO or after ALF).
 - Performance comparison of residual scaling on and off.

Test	# Param	# Hidden layers(N)	# Models (X)	Position	Additional info.
1	~1M*4	12	4	Prior SAO	Use QP as input
2	~0.5M*4	6	4	Prior SAO	Use QP as input
3	~1M*4	Subtest #1 without residual scaling			
4	~1M*12	12	12	Prior SAO	
5	~1M*12	12	12	After ALF	

Simulation results

- QP: 22,27,32,37,42
- Anchor: VTM-10.0 GOP size = 32

		Y	U	V	EncT(%)	DecT(x)	Note
1	RA	-5.32%	-13.15%	-12.95%	140%	149x	
	AI	-4.83%	-10.09%	-10.65%	125%	85x	
2	RA	-4.57%	-11.77%	-11.62%	122%	78x	Half sized model
	AI	-4.20%	-8.77%	-9.92%	114%	45x	
3	RA	-4.78%	-11.71%	-11.94%	139%	148x	Residual scaling off
	AI	-4.60%	-9.42%	-10.36%	123%	85x	
4	RA	-4.95%	-12.73%	-12.58%	208%	148x	12 models
	AI	-4.39%	-9.94%	-10.67%	163%	85x	
5	RA	-4.02%	-12.08%	-12.25%	208%	148x	12 models, after ALF
	AI	-3.26%	-8.28%	-9.28%	163%	85x	

Thank you!

