

# JVET-V0114

## EE1-1.3: Test on Neural Network-based In-Loop Filter with No Deblocking Filtering Stage

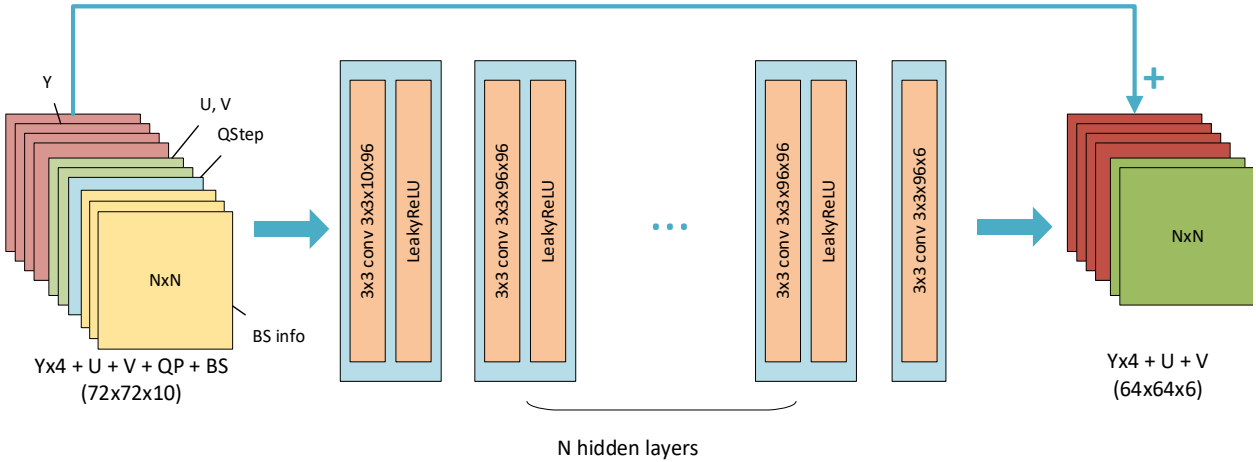
H. Wang, J. Chen, A. M. Kotra, K. Reuzé, M. Karczewicz (Qualcomm)



# Network Structure

EE test of JVET-U0115

- Input: a 144x144 YUV420 block
  - The input block is converted into a 72x72x6 tensor
  - QP info is used as an additional input plane
  - Deblocking boundary strength (BS) info is used as input planes
  - Pixel values are normalized to [0.0,1.0]
- Output: Residues to be added to input samples
- Number of convolutional layers: 14
- Number of feature maps: 96
- Number of parameters: ~1M



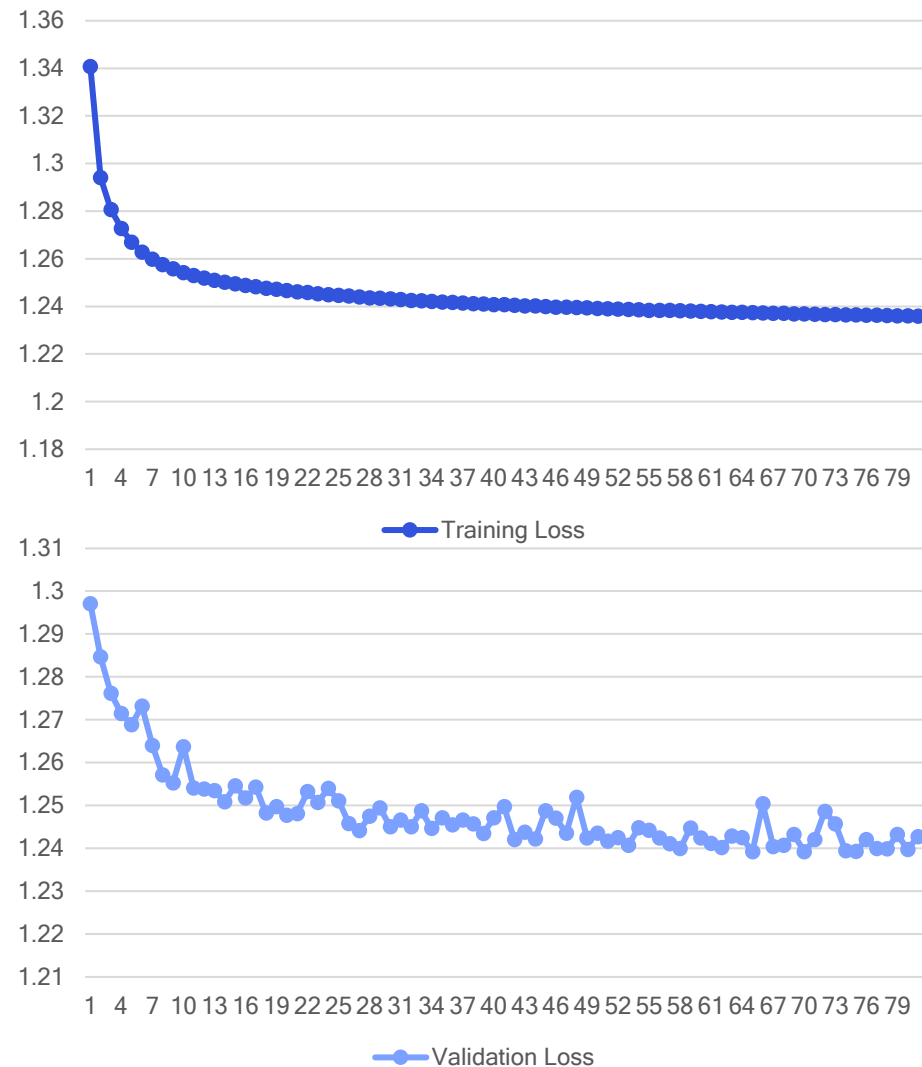
Network Information in Inference Stage		
Mandatory	GPU Type	CPU Only
	Framework:	Tensorflow v2.2.0
	Number of GPUs per Task	0
	Total Parameter Number	1,010,406*4Models
	Parameter Precision (Bits)	32
	Memory Parameter (MB)	3.85 * 4 models
	MAC (Giga)	3.19e-4 Giga/pixel
Optional	Total Conv. Layers	14
	Total FC Layers	0
	Total Memory (MB)	
	Batch size:	1
	Patch size	72x72x10
	Peak Memory Usage	1.9MB(inputsize 144x144)

# Training process

- Training data set: BVI-DVC, DIV2K
- Training data generation
  - QPs: 20, 25, 30, 35, 40, 45
  - Configurations used to generate training data:
    - AI, RA
- Number of trained models:
  - 4 ( 2 Intra picture based + 2 Inter picture based)

Network Information in Training Stage		
Mandatory	GPU Type	GPU: NVIDIA Tesla V100-SXM2-32GB
	Framework:	Tensorflow 2.2.0
	Number of GPUs per Task	1
	Epoch:	~80
	Batch size:	82Kx16
	Training time:	106h
	Training data information:	BVI-DVC, DIV2K
Optional	Training configurations for generating compressed training data (if different to VTM CTC):	QP=20,25,30,35,40,45
	Number of iterations	
	Patch size	
	Learning rate:	2.00E-04
	Optimizer:	ADAM
	Loss function:	Weighted L2
	Preprocessing:	Convert 144x144 YUV420 signal to 6 72x72 blocks. Normalize to 0~1

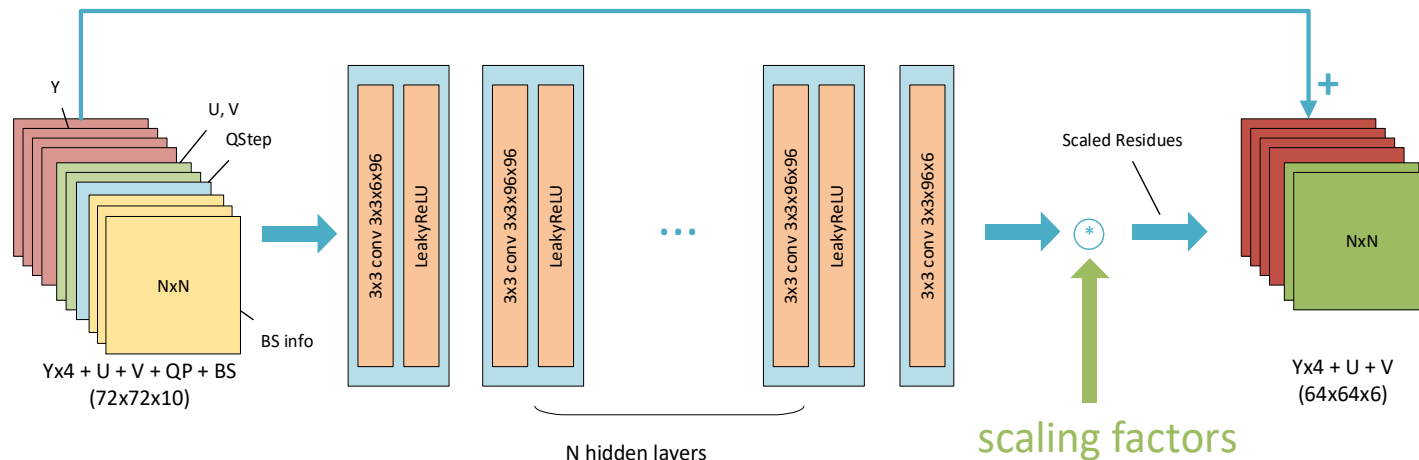
# Training process



Network Information in Training Stage		
Mandatory	GPU Type	GPU: NVIDIA Tesla V100-SXM2-32GB
	Framework:	Tensorflow 2.2.0
	Number of GPUs per Task	1
	Epoch:	~80
	Batch size:	82Kx16
	Training time:	106h
	Training data information:	BVI-DVC, DIV2K
	Training configurations for generating compressed training data (if different to VTM CTC):	QP=20,25,30,35,40,45
Optional	Number of iterations	
	Patch size	
	Learning rate:	2.00E-04
	Optimizer:	ADAM
	Loss function:	Weighted L2
	Preprocessing:	Convert 144x144 YUV420 signal to 6 72x72 blocks. Normalize to 0~1

# VTM integration

- Integrated into VTM-11.0
- For each frame, one model is selected from the 4 candidates
- Slice/CTU level on/off is supported
- A scaling factor is signaled for each color component in the picture header



- Deblocking filter module is used to generate  $BS$  information but actual filtering is not applied

# Simulation results

- QP: 22,27,32,37,42
- Anchor: VTM-11.0

EE1-1.3 test results

Random access Main10						
BD-rate Over VTM-11.0						
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT	bit DIFF
Class A1	-7.38%	-9.22%	-10.40%	150%	13659%	0%
Class A2	-7.35%	-13.75%	-10.84%	146%	12460%	1%
Class B	-6.90%	-13.95%	-13.85%	150%	14461%	1%
Class C	-7.49%	-16.12%	-16.00%	135%	13931%	1%
Class E						
<b>Overall</b>	-7.24%	-13.54%	-13.13%	145%	13740%	1%
Class D	-8.90%	-17.84%	-17.72%	135%	14900%	1%
Class F	-3.67%	-11.26%	-10.68%	176%	18533%	0%

All Intra Main10						
BD-rate Over VTM-11.0						
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT	bit DIFF
Class A1	-6.31%	-9.77%	-11.95%	152%	9419%	0%
Class A2	-6.00%	-12.55%	-11.10%	130%	7238%	0%
Class B	-6.06%	-11.78%	-13.29%	133%	8020%	0%
Class C	-6.61%	-12.27%	-14.36%	121%	6220%	0%
Class E	-8.66%	-11.70%	-12.68%	133%	9100%	0%
<b>Overall</b>	-6.64%	-11.67%	-12.84%	133%	7916%	0%
Class D	-6.80%	-12.21%	-14.66%	112%	5943%	1%
Class F	-4.60%	-10.42%	-10.74%	109%	6202%	0%

Supplemental results of refined training

Random access Main10						
BD-rate Over VTM-11.0						
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT	bit DIFF
Class A1	-7.51%	-9.79%	-12.56%	157%	14878%	0%
Class A2	-7.62%	-13.69%	-11.80%	153%	13608%	1%
Class B	-7.09%	-15.28%	-15.74%	151%	15220%	1%
Class C	-7.75%	-17.39%	-17.79%	137%	14333%	1%
Class E						
<b>Overall</b>	-7.46%	-14.43%	-14.86%	148%	14580%	1%
Class D	-9.11%	-18.93%	-19.93%	136%	15146%	1%
Class F	-3.84%	-12.17%	-11.60%	177%	19039%	0%

All Intra Main10						
BD-rate Over VTM-11.0						
	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT	bit DIFF
Class A1	-6.24%	-11.01%	-13.16%	155%	9951%	0%
Class A2	-6.10%	-12.94%	-12.09%	137%	8130%	0%
Class B	-6.16%	-12.50%	-14.47%	134%	8432%	0%
Class C	-6.79%	-13.67%	-16.37%	119%	6296%	1%
Class E	-8.80%	-13.19%	-14.87%	130%	9186%	0%
<b>Overall</b>	-6.74%	-12.70%	-14.35%	134%	8254%	0%
Class D	-6.92%	-13.15%	-17.23%	117%	6472%	1%
Class F	-4.83%	-12.50%	-12.11%	115%	7398%	0%

Thank you!

