|  |
| --- |
| **INTERNATIONAL ORGANIZATION FOR STANDARDIZATION ORGANISATION INTERNATIONALE DE NORMALISATION ISO/IEC JTC 1/SC 29/WG 5 MPEG JOINT VIDEO CODING TEAM WITH ITU-T SG 16** |
| **ISO/IEC JTC 1 / SC 29 / WG 5 N 164** |
| **Mainz – 20–28 October 2022** |
| |  |  | | --- | --- | | **Title:** | **Exploration experiment on neural network-based video coding (EE1)** | | **Source:** | **Convenor (Jens-Rainer Ohm)** | | **Type:** | **General** | | **Subtype:** | **Other** | | **Status:** | **Approved** | | **Date:** | **2022-11-14** | | **Expected Action:** | **Info** | | **Action due date:** | **N/A** | | **No. of pages** | **14** (without this cover page) | | **Email of convenor:** | **ohm @ ient . rwth-aachen . de** | | **Committee URL:** | **https://sd.iso.org/documents/ui/#!/browse/iso/iso-iec-jtc-1/iso-iec-jtc-1-sc-29/iso-iec-jtc-1-sc-29-wg-5** | |

|  |  |
| --- | --- |
| **Joint Video Experts Team (JVET)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29**  28th Meeting, Mainz, DE, 20–28 October 2022 | Document: JVET-AB2023-v2 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **EE1: Summary of Exploration Experiments on Neural Network-based Video Coding** | | |
| *Status:* | Output document to JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | E. Alshina, F. Galpin, Y. Li, M. Santamaria, J. [Ström](mailto:jacob.strom@ericsson.com), H. Wang, L. Wang and Z. Xie | Tel: Email: | [elena.alshina@huawei.com](about:blank)  [franck.galpin@interdigital.com](mailto:franck.galpin@interdigital.com)  [yue.li@bytedance.com](mailto:yue.li@bytedance.com)  [maria.santamaria\_gomez@nokia.com](mailto:maria.santamaria_gomez@nokia.com)  [hongtaow@qti.qualcomm.com](about:blank)  [jacob.strom@ericsson.com](mailto:jacob.strom@ericsson.com)  [liqiangwang@tencent.com](mailto:liqiangwang@tencent.com)  [xiezhihuang@oppo.com](mailto:xiezhihuang@oppo.com) |
| *Source:* | EE Coordinators | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document summarizes Exploration Experiment 1 (EE1) tests to be performed between the JVET-AB and JVET-AC meetings to evaluate **Neural Network-based Video Coding (**NNVC) technologies, analyze their performance, and analyze their complexity aspects.

# Introduction

Group continues evaluation of new promising NN-based video coding technologies, answering questions and addressing suggestions from JVET members made during presentation of NN-based technologies at JVET-AB meeting.

NNVC common SW base (***NNVC***) and training scripts are highly encouraged to be used by all EE1 proponents, but not mandated for this round of EE1.

The most promising technologies recommended by JVET **will undergo procedure of cross-check for the training**.

Tests will be conducted in three categories: enhancement filters (both post- and in-loop), super-resolution and NN-based Intra coding.

All proponents **must** report results relatively to the default configuration of AhG11 anchor NNVC-3.0 in <https://vcgit.hhi.fraunhofer.de/jvet-ahg-nnvc/VVCSoftware_VTM> [[[1]](#endnote-2)] and the reported template recommended by AhG11 (please use up-to-date version). Both existing neural network-based filters are disabled for the the default configuration of AhG11 anchor NNVC-3.0.

If test is improving or modifying ***NNVC*** ***filter set #0*** or ***NNVC*** ***filter set #1***, then results relatively to corresponding filter in NNVC-3.0 are requested.

Proponents in Super Resolution category are requested to report results both relatively to AhG11 anchor and also relatively to adaptive resolution coding with RPR filter (Test 2.1). For super-resolution category, proponents are recommended to provide results with 10% rate matching or with rate points that provide close enough quality matching to the anchor. It is requested to provide RD-curves for all six 4K sequences in comparison with AhG11 anchor.

Proponents are encouraged to report both CPU and GPU decoding run time.

Test results and complexity analysis reporting template [1] are expected to be uploaded together with final software by the T4 deadline specified in section “Timeline”.

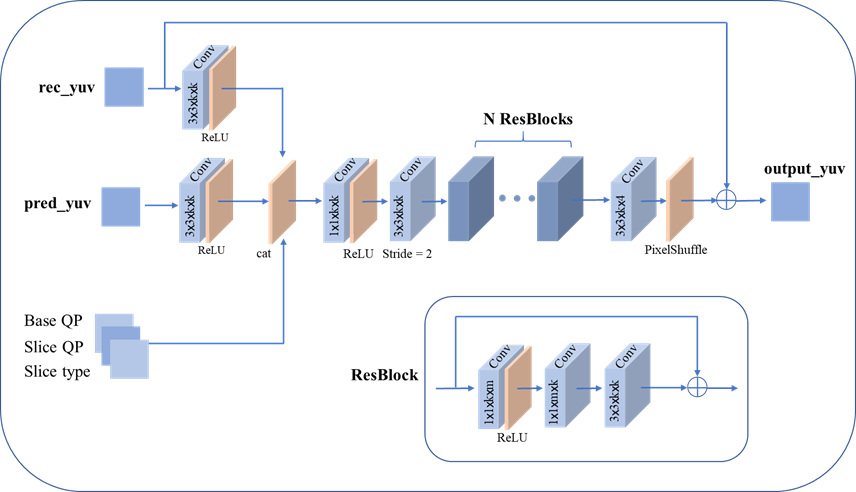
Discussions with regards to this EE are expected to be conducted in JVET reflector.

# Exploration experiments on Enhancement filters

## NNVC filter set#0 based

[JVET-AB0083](https://jvet-experts.org/doc_end_user/current_document.php?id=12010) EE1-1.8: More refinements on NN based in-loop filter with a single model [L. Wang, X. Xu, S. Liu (Tencent), Z. Xie, Y. Yu, H. Yu, D. Wang (OPPO)]

This test evaluates the fine-tuned model for the filter set #0. The original model is trained based on the training dataset generated by the codec with EncDbOpt disable. However, EncDbOpt was enabled in the common test conditions and the performance is largely improved by about 0.8% under RA configuration, so it is proposed to fine-tune the model based on the training dataset generated by the latest version. The other implementation details are identical to those in the filter set #0.

**

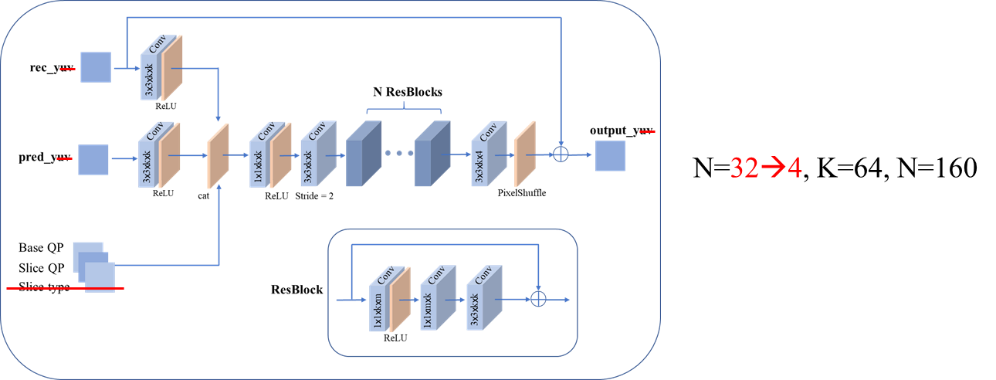
Tests to be studied in EE:

**Test 1.1.1**: Test the fine-tuned model for the filter set #0.

**Test 1.1.2**: Optimize the trade-off by slightly modifying the network design or the implementation.

[JVET-AB0146](https://jvet-experts.org/doc_end_user/current_document.php?id=12073) EE1-1.8-related: encoder-only optimization for NN based in-loop filter with a single model [L. Wang, X. Xu, S. Liu (Tencent)]

This test evaluates the encoder-only filter models in the partitioning decision process for the filter set #0. The proposed network structure is similar to the one in the filter set #0, but there are several differences. Firstly, only four Resblocks are used in network structure of the proposed filter. Secondly, there are two models for I slices and B slices, respectively. Therefore, the slice-type information is removed from the inputs. Thirdly, only luma related images are fed into the network.

**

Tests to be studied in EE:

**Test 1.2**: Test the encoder-only filter models for filter set #0.

It was requested by JVET the ***training scripts*** for encoder only model to be provided.

[JVET-AB0158](https://jvet-experts.org/doc_end_user/current_document.php?id=12085) [AHG11] On chroma order adjustment in NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

A chroma order adjustment method based on JVET-AA0088 was proposed to improve the coding performance of *NNVC* filter set #0 in this contribution. A neural network-based in-loop filter with constrained memory size and low complexity was presented in JVET-AA0088 (shown in Fig. 1) and was adopted as filter set #0 in *NNVC*, in which the chroma information including U and V components are part of the inputs/outputs to the neural network. JVET-AB0158 proposed to allow the input/output order switch between the U and V components at the frame level in the inference stage. No other change was proposed in this contribution and the network architecture kept the same as the one in JVET-AA0088.

A frame-level flag is signaled to indicate whether the chroma order adjustment method is used.

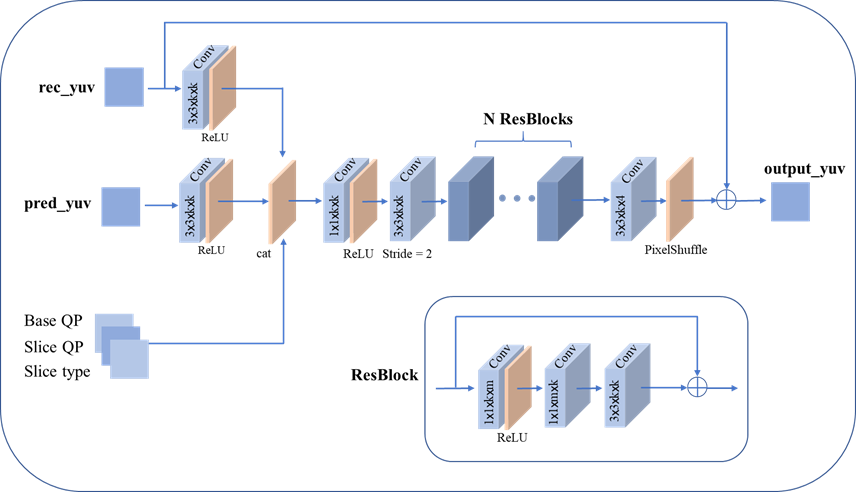


Fig. 1: Network architecture of JVET-AA0088

Tests to be studied in EE:

**Test** **1.3** Optimize and test JVET-AB0158 as proposed.

[JVET-AB0159](https://jvet-experts.org/doc_end_user/current_document.php?id=12086) [AHG11] On adjustment of residual for NNLF [Z. Dai, Y. Yu, H. Yu, D. Wang (OPPO)]

An adjustment of residual for NNLF’s output based on JVET-AA0088 was proposed to improve the coding performance of *NNVC* filter set #0 in this contribution. A neural network-based in-loop filter with constrained memory size and low complexity was presented in JVET-AA0088 (shown in Fig. 2) and was adopted as filter set #0 in *NNVC*. JVET-AB0159 proposed that a frame level residual offset can be used to adjust the output of the NNLF in the inference stage. More specifically, the residual is adjusted by reducing the magnitude of the residual at each pixel by a small offset value, and the offset candidates are {1, 2}. No other change was proposed in this contribution and the network architecture kept the same as the one in JVET-AA0088.

Two frame level flags are signalled to indicate whether the residual offset method is enabled for luma and chroma, respectively. If residual offset is enabled, another frame level flag is signalled to indicate what offset value is used for adjusting the corresponding residual colour component(s).

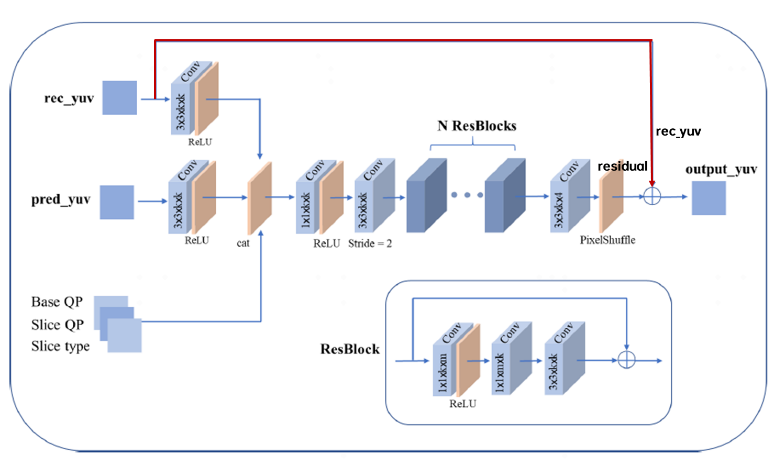


Fig. 2: Network architecture of JVET-AA0088

Tests to be studied in EE:

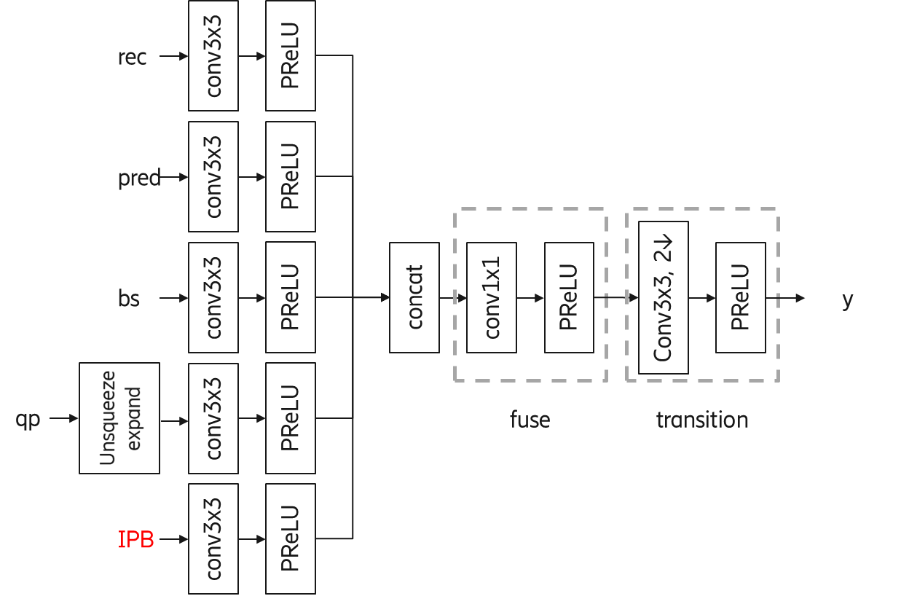
**Test** **1.4.1** Optimize and test JVET-AB0159 as proposed (w/o retraining).

**Test 1.4.2** Test the combination of JVET-AB0158 and JVET-AB0159 on top of both *NNVC* loop filter set #0 and set #1.

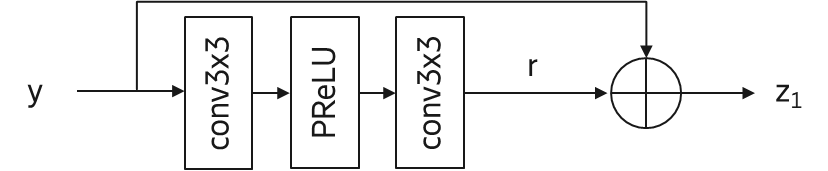
## NNVC filter set#1 based

[JVET-AB0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11967) EE1-1.5: One luma model with IPB and/or skip for filtering intra and inter luma slices [D. Liu, J. Ström, M. Damghanian, P. Wennersten, K. Andersson (Ericsson)]

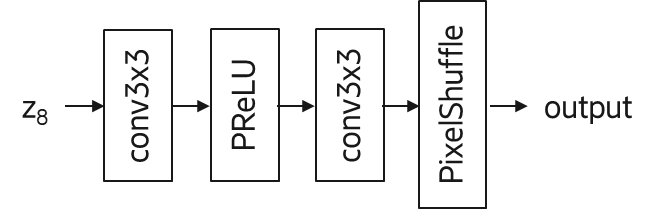
JVET-AB0052 describes using a single luma model for both Intra and Inter slices with an extra IPB input (block type I/P/B) on top of filter set #1. To align with the NN structure from JVET-AB0053 to be included in NNVC-3.0 (*NNVC* -1.0) which does not have partition and attention modules for the intra luma model, the attention modules for the testing IPB model are removed. The ***training is to be cross-checked***. It is also proposed at the meeting to extend the IPB idea to the chroma model.



*Figure 1. Head of network. The inputs are combined to form the input y to the next part of the network.*



*Figure 2. The k-th residual block (k=0..7). The output y of the head is fed into a first residual block with input z0=y, which also takes the inputs rec, pred, part, bs, qp, and IPB. The output z1 is then fed into another such residual block.*



*Figure 3. The output of the last residual block is fed into this last part of the network.*

Tests to be studied in EE:

**Test** **1.5.1** Training cross-check for the single luma model for both inter luma and intra luma slices in filter set #1, with the IPB input, without the attention modules.

**Test 1.5.2** Extending the IPB usage to the chroma model, i.e., having one chroma model for inter chroma and intra chroma slices.

[JVET-AB0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11968) EE1-1.2: NN intra model without attention and partitioning strength [J. Ström, D. Liu, K. Andersson, P. Wennersten, M. Damghanian, R. Yu (Ericsson)]

In JVET-AB0053, it was demonstrated that the network architecture in *NNVC* 1.0 filter set #1 could be simplified by removing the partitioning input and removing the attention branch for the intra luma filter. It was proposed at the meeting to also try removing partitioning as an input for the intra chroma filter. There is no need to remove the attention from the intra chroma filter since this is not available. Therefore, the following tests are proposed:

**Test** **1.6.1** Intra chroma filter with the same architecture as *NNVC* 1.0 filter set 1 except that partitioning input has been removed, retrained.

**Test 1.6.2** Intra chroma filter with the same architecture as *NNVC* 1.0 filter set 1, retrained from scratch using the same methodology as in EE1-1.2.1.

[JVET-AB0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11988) EE1-1.4: Deep In-Loop Filter with Additional Input Information [Y. Li, K. Zhang, L. Zhang (Bytedance)]

This test investigates the neural network-based in-loop filtering method proposed in JVET-AB0073. Two aspects are proposed to improve the CNN models. Aspect #1: Addition input information can be input to the in-loop filtering network. Aspect #2: The input samples of CNN can be flipped and the output samples of CNN can be flipped back.The network backbone is based on residual blocks. Other design elements including parameter selection, residual scaling, combination of NN filtering and deblocking filtering, etc. remain the same as the filter set #1 in NNVC common software base.



Tests to be studied in EE:

**Test 1.7.1** Optimize and test the method as in JVET-AB0073.

**Test 1.7.2** Optimize and test the aspect #1 in JVET-AB0073, i.e. feeding additional inputs into the network.

**Test 1.7.3** Optimize and test the aspect #2 in JVET-AB0073, i.e. flipping the input samples and flipping back the output samples.

[JVET-AB0141](https://jvet-experts.org/doc_end_user/current_document.php?id=12068) EE1-related: QP-based loss function design for NN-based in-loop filter [C. Zhou, Z. Lv, J. Zhang (vivo), W. Chen, J. Guo, B. Ai (BJTU)]

In this test, a new loss function is defined where quantization parameter is used as a factor for the loss calculation in the training stage. In NN-based in-loop filter, the commonly used loss functions, L1 and L2 have not considered the information related to video coding. Since the training samples with different quantization parameters have different training difficulties, the new designed loss function can be formulated as follows:

or

where denotes the weight assigned to the conventional loss function according to the quantization parameters.

Tests to be studied in EE:

**Test 1.8.1** Test the method as in JVET-AB0141 on top of *NNVC* filter set #1.

**Test 1.8.2** Optimize the loss function calculation.

## Not in NNVC NN-filters

[JVET-AB0164](https://jvet-experts.org/doc_end_user/current_document.php?id=12091) EE1-1.7: Capacity Ablation of CNN-based in-loop filtering [S. Eadie, H. Wang, M. Coban, M. Karczewicz (Qualcomm)]

This test studies the model originally proposed in JVET-AA0131. The complexity-performance trade-off of is analysed independently for its luma and chroma models based on the channels’ complexities and subsequent required processing capacities. Specifically, the number of residual blocks is ablated across {16, 24, 32} with n=32 representing the original model from AA0131.



Tests to be studied in EE:

**Test** **1.9.1** Optimize the floating-point/fixed-point SADL implementation of the models.

**Test 1.9.2** Cross check the **training process**

**Test 1.9.3** Optimize the trade-off between performance and complexity by modifying the network structure (e.g. number of res blocks, input information, ResBlock structure etc).

**Test 1.9.4** Optimize the trade-off between performance and complexity by reducing the total number of models

**Test 1.9.5** A combination of Test 1.9.3 and Test 1.9.4

[JVET-AB0136](https://jvet-experts.org/doc_end_user/current_document.php?id=12063) AHG11: Complexity Reduction on Neural-Network Loop Filter [J. N. Shingala, S. Kadaramandalgi, A. Shyam (Ittiam), T. Shao, A. Arora, P. Yin, Sean McCarthy (Dolby)]

This test investigates the complexity reduction techniques on neural network-based loop filter proposed in JVET-AB0136. It’s proposed to use CP decomposition plus fusing adjacent 1x1 convolution to reduce the complexity of NNLF model. Additionally, it’s also proposed to further split architecture for luma and chroma components between input network and output network.

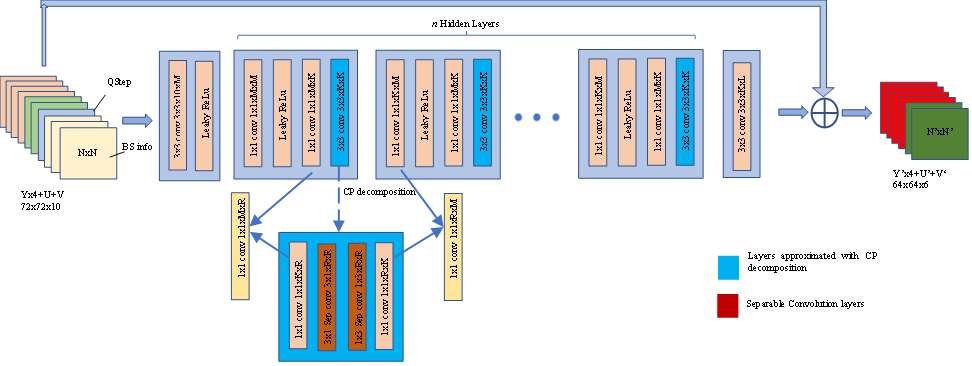


Figure 1 CP Decomposition + fusion of 1x1 conv layers of JVET-X0140 Baseline Model

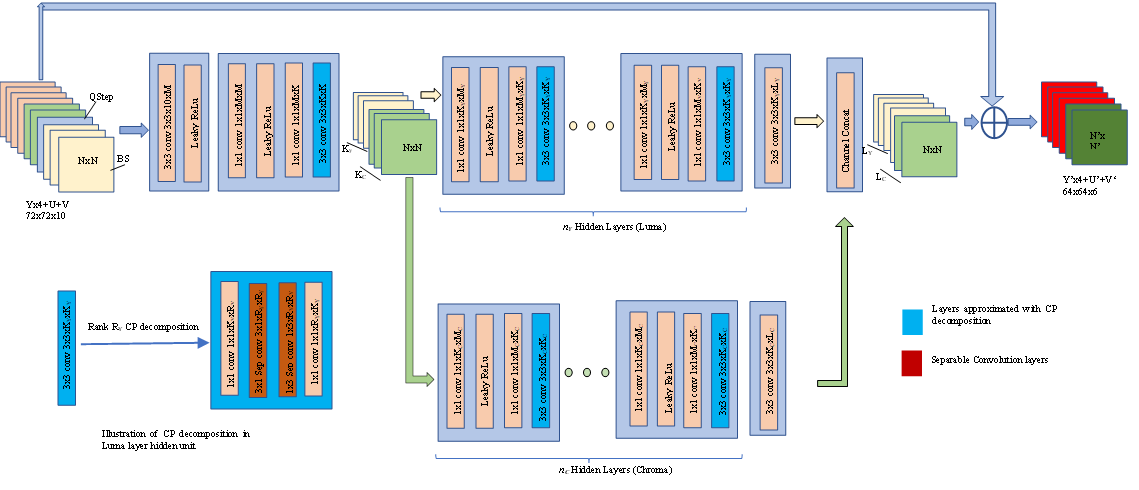


Figure 2 Split luma and chroma model + CP Decomposition + fusion of 1x1 conv layer

Tests to be studied in EE:

**Test 1.10.1** Test the model with CP decomposition and fusing adjacent 1x1 convolution with following model parameters:

* Number of hidden layers: n = 11
* Feature maps and rank: K = 24, R = 24, M = 72
* Model complexity = 16.2 kMAC/pixel

**Test** **1.10.2** Test the model with CP decomposition, fusing adjacent 1x1 convolution and split architecture for luma and chroma components with following model parameters:

* Number of hidden layers for luma and chroma: nY = nC = 10
* Feature maps and rank for luma and chroma split (24L, 8C):
  + Luma: KY = 24, RY = 24, MY = 72
  + Chroma: KC = 8, RC = 8, MC = 24
* Model complexity: 17.7 kMAC/pixel

Both the above tests to be verified for CPU inference as follows:

* NNVC SW: AI and RA inference in current NNVC-3.0 software (floating point).
* SADL: Fixed point integer implementation using SADL library to validate matching filtered outputs of encoder-decoder. Assess the impact on coding performance due to integer quantization.

## NN-based post-filters

[JVET-AB0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11963) EE1-1.1: Content-adaptive post-filter with SADL inference and signalling of NN post-filter characteristics and activation SEI messages [M. Santamaria, R. Yang, F. Cricri, J. Lainema, H. Zhang, R. G. Youvalari, M. M. Hannuksela (Nokia)]

This test studies an NN content-adaptive post-filter. The inputs to the filter are the reconstructed luma samples, chroma samples and frame QP.



Tests to be studied in EE:

**Test** **1.11.1** **Training cross-check**. Float32 inference.

**Test 1.11.2** Int16 inference.

# Super-resolution NN-based post-filters

[JVET-AB0102](https://jvet-experts.org/doc_end_user/current_document.php?id=12029) AHG11/EE1-related: Updates on RPR encoder and filters [J. Nam, S. Yoo, J. Lim, S. Kim (LGE)]

This contribution proposes an RPR encoder with four available scale factors {x2.0, x1.5, x1.25, x1.0}. The scale factor is selected at GOP level based on the initial QP and PSNR value of the first picture in the GOP and it is applied to all pictures within the GOP. This contribution also suggests to increase RPR filter and post filter taps.

Tests to be studied in EE:

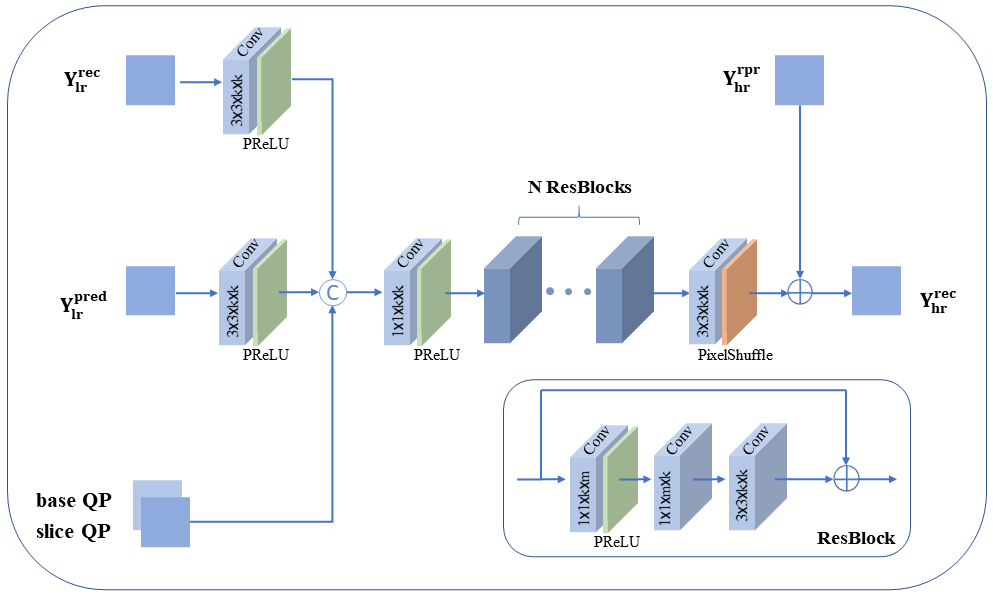
**Test 2.1.1** Test adaptive resolution selection with scale ratio 1.0 and 2.0

**Test 2.1.2** Test adaptive resolution selection with scale ratio 1.0, 1.25, 1.5 and 2.0

**Test 2.1.3** Test increasing post-filter taps on top of Test 2.1.2

[JVET-AB0098](https://jvet-experts.org/doc_end_user/current_document.php?id=12025) EE1-2.3 related: GOP Level Adaptive Resampling with CNN-based Super Resolution [R. Chang, L. Wang, X. Xu, S. Liu (Tencent)]

This test studies the GOP level adaptive resampling method with CNN-based super resolution proposed in JVET-AB0098. At the GOP level, the proposed method can adaptively select a scale factor from ×1.0 (original size) and ×2.0 (half size) to determine the encoding resolution, and the designed CNN-based super-resolution will be used for the half size.

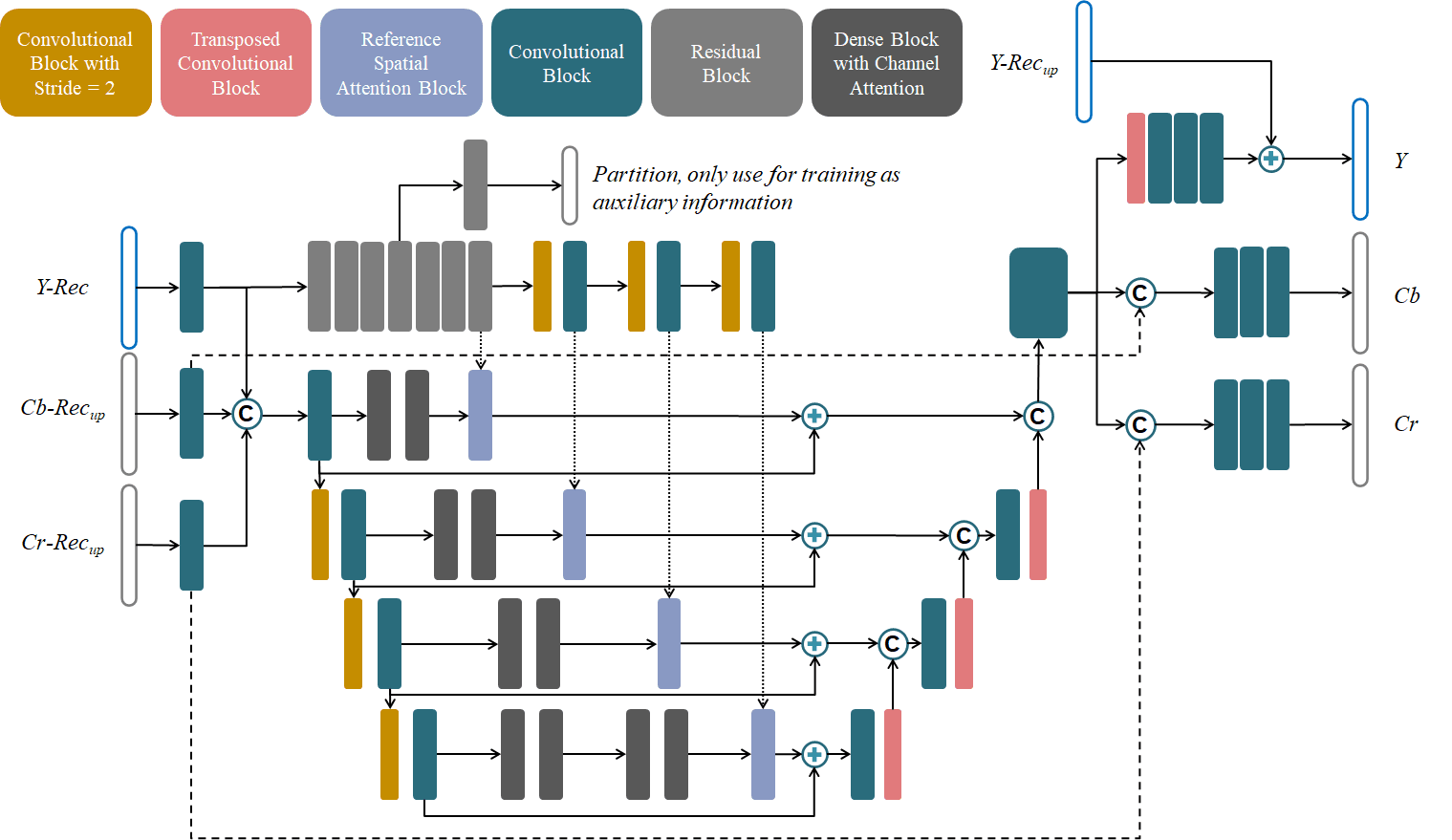


Tests to be studied in EE:

**Test 2.2.1** Test this method with 10% rate matching.

**Test 2.2.2** Optimize the trade-off between performance and complexity.

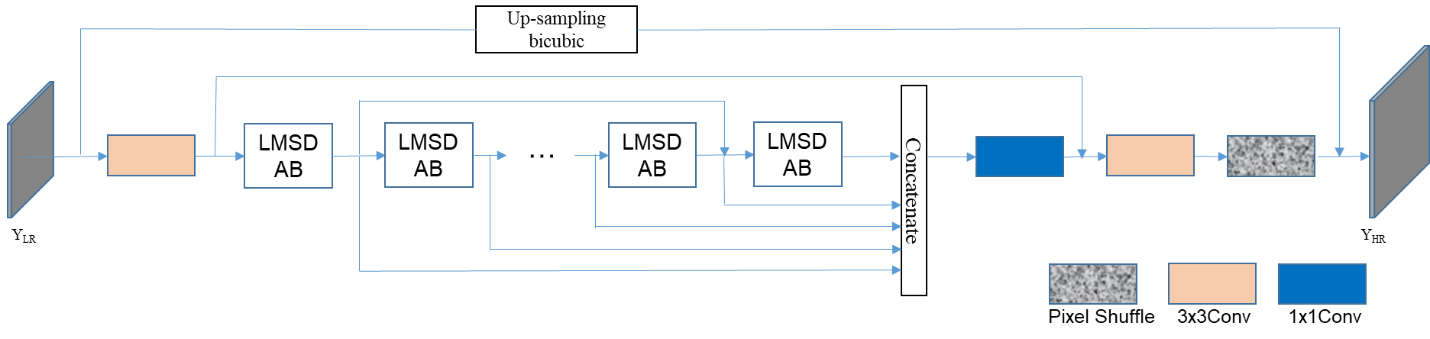
[JVET-AB0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11991) EE1-2.1: RPR-Based Super-Resolution Guided by Partition Information [Q. Han, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]

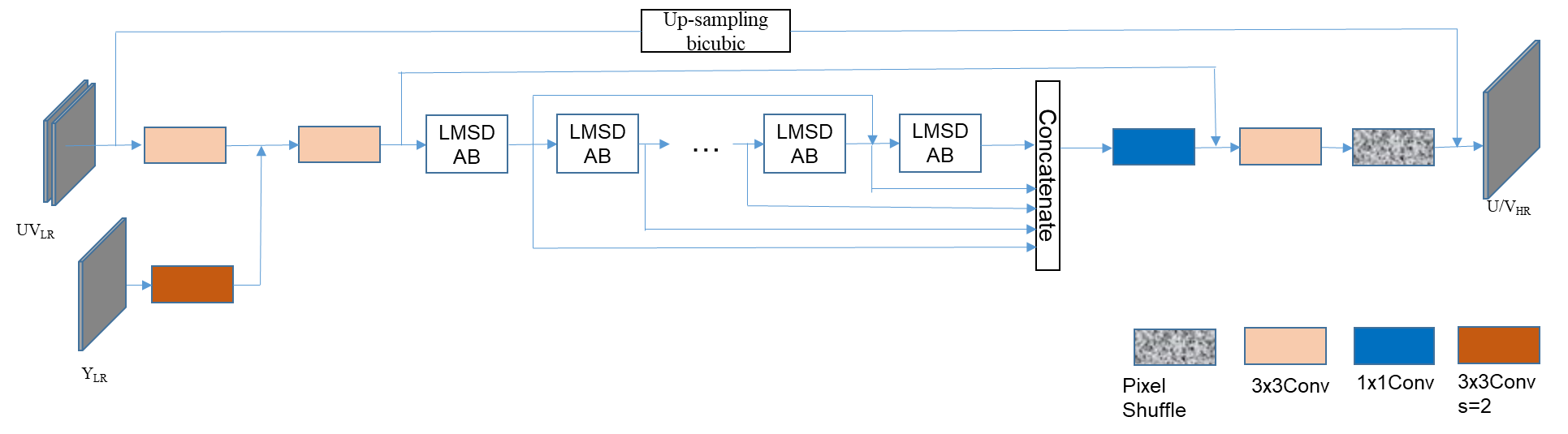


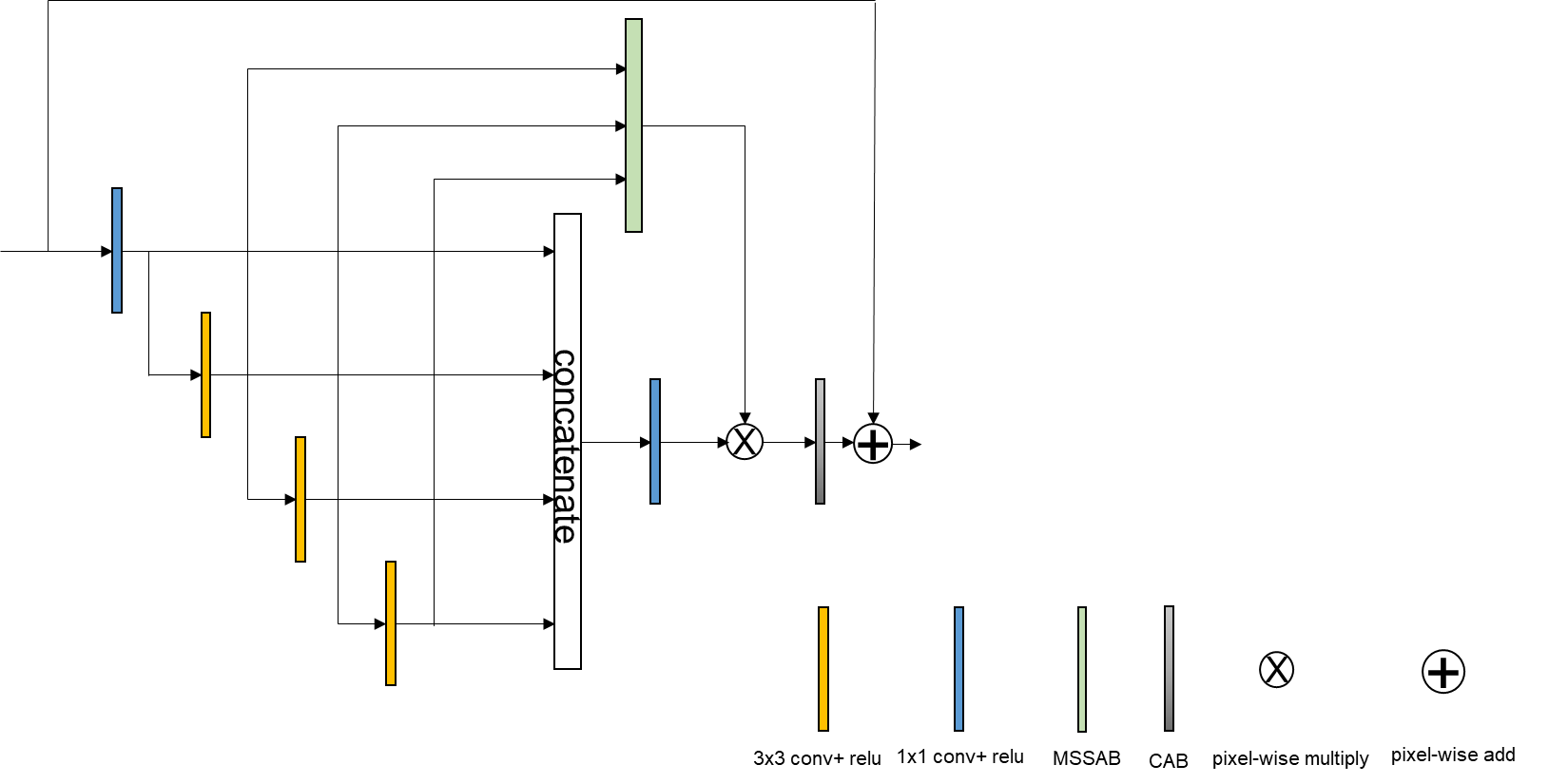
**Test 2.3.1** Introduce QP as input to merge the five models into one to reduce the memory consumption and solve the QP mismatch problem.

**Test 2.3.2** Combine the SR filter with JVET-Z0065 in RPR encoder for adaptive resolution selection to match the 10% rate.

[JVET-AB0093](https://jvet-experts.org/doc_end_user/current_document.php?id=12020) EE1-2.2 related: Lightweight CNN Filter for Super-Resolution with RPR functionality in VVC [S. Huang, C. Jung (Xidian Univ.), Y. Liu, M. Li (OPPO)]







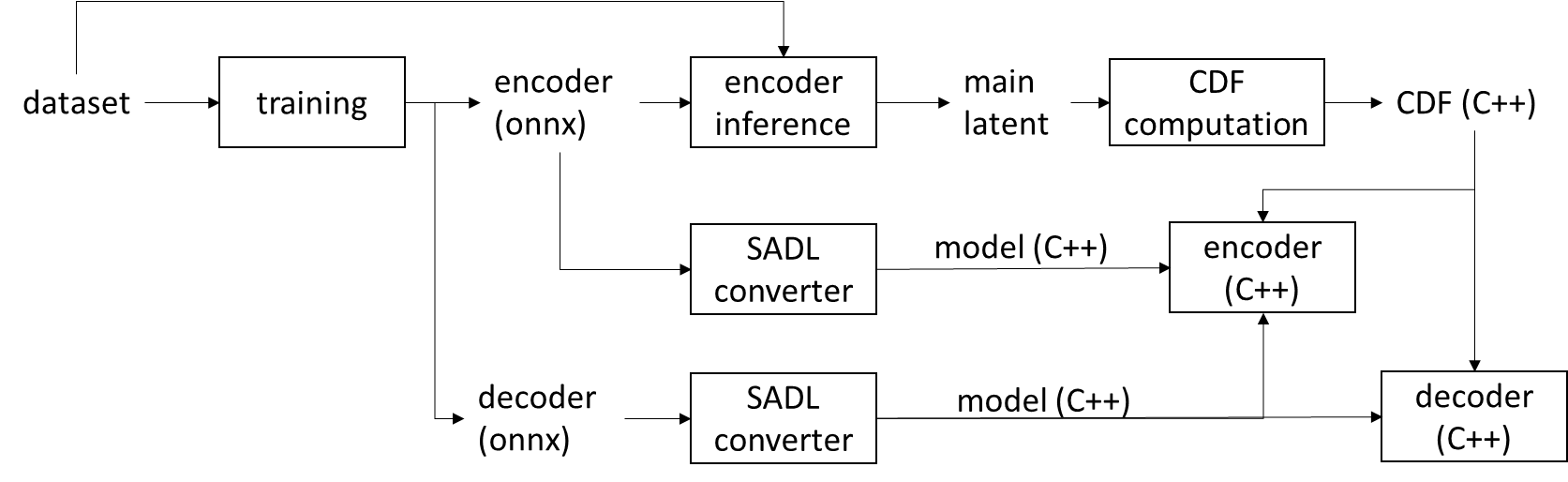
Tests to be studied in EE:

**Test 2.4.1** Test the network architecture as proposed in JVET-AB0093.

**Test 2.4.2** Combine JVET-AB0093 (LMSDANet) with JVET-Z0065 to provide adaptive resolution selection for matching 10% rate.

# NN-Intra coding

[JVET-AB0125](https://jvet-experts.org/doc_end_user/current_document.php?id=12052) AHG11 - CompressAI models integration using SADL [F. Galpin, F. Levebvre, F. Racapé (InterDigital)]



**Test 3.1** The test for JVET-AB0125 consists in testing the performance of the produced stand-alone encoder/decoder after integerization of the model. Comparison with float model will be performed on the Kodak dataset to assess the impact of the integerization. Models produced using the procedure described in CompressAI repository will be provided for crosscheck. Comparison of inference on different platforms and/or compilers will be disclosed.

[JVET-AB0149](https://jvet-experts.org/doc_end_user/current_document.php?id=12076) Non-EE1: neural network-based intra prediction with learned mapping to VVC intra prediction modes [T. Dumas, F. Galpin, P. Bordes (InterDigital)]

The EE1 test for JVET-AB0149 consists of cross-checking the training stage.

* using the provided training scripts to re-train the neural network-based intra prediction mode presented in JVET-AB0149. This neural network-based intra prediction mode comprises 7 different fully-connected neural networks.
* **Test 3.2.1** running the provided modified version of *NNVC* -1.0, Filter-Set-0 being activated, Filter-Set-1 being deactivated, and the neural network-based intra prediction mode being activated, to reproduce approximately the BD-rate gains reported in JVET-AB0149.
* **Test 3.2.2** running the provided modified version of *NNVC* -1.0, Filter-Set-1 being activated, Filter-Set-0 being deactivated, and the neural network-based intra prediction mode being activated, to confirm that the BD-rate gains brought by the neural network-based intra prediction are almost equivalent when this intra mode is combined with either Filter-Set-0 or Filter-Set-1.

# Cross-check allocation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test | proposal | tester | Cross-checker | Comment |
| NNVC filter set #0 based | | | | |
| [Test 1.1](#TEST1_1) | [JVET-AB0083](https://jvet-experts.org/doc_end_user/current_document.php?id=12010) | [L. Wang](mailto:liqiangwang@tencent.com) | [Z. Xie](mailto:xiezhihuang@oppo.com)  (OPPO) | **Training** and inference |
| [Test 1.2](#TEST1_2) | [JVET-AB0146](https://jvet-experts.org/doc_end_user/current_document.php?id=12073) | [L. Wang](mailto:liqiangwang@tencent.com) | [Z. Xie](mailto:xiezhihuang@oppo.com)  (OPPO) | Only inference |
| [Test 1.3](#TEST1_3) | [JVET-AB0158](https://jvet-experts.org/doc_end_user/current_document.php?id=12085) | [Z. Dai](mailto:daizhenyu@oppo.com) | [L. Wang](mailto:liqiangwang@tencent.com)  (Tencent) | Only inference |
| [Test 1.4](#TEST1_4) | [JVET-AB0159](https://jvet-experts.org/doc_end_user/current_document.php?id=12086) | [Z. Dai](mailto:daizhenyu@oppo.com) | [L. Wang](mailto:liqiangwang@tencent.com)  (Tencent) | Only inference |
| NNVC filter set #1 based | | | | |
| [Test 1.5](#TEST1_5) | [JVET-AB0052](https://jvet-experts.org/doc_end_user/current_document.php?id=11967) | [D. Liu](mailto:du.liu@ericsson.com) | [R. Chang](mailto:renjiechang@tencent.com)  (Tencent) | **Training** and inference |
| [Test 1.6](#TEST1_6) | [JVET-AB0053](https://jvet-experts.org/doc_end_user/current_document.php?id=11968) | [J. Ström](mailto:jacob.strom@ericsson.com) | M. Santamaria (Nokia) | **Training** and inference (SADL integer) |
| [Test 1.7](#TEST1_7) | [JVET-AB0073](https://jvet-experts.org/doc_end_user/current_document.php?id=11988) | [Y. Li](mailto:yue.li@bytedance.com) | T. Shao (Dolby)  C. Zhou (Vivo) | **Training** and inference |
| [Test 1.8](#TEST1_8) | [JVET-AB0141](https://jvet-experts.org/doc_end_user/current_document.php?id=12068) | [C. Zhou](mailto:chuan.zhou@vivo.com) | Y. Li (Bytedance) | **Training** and inference |
| Not in NNVC NN-filters | | | | |
| [Test 1.9](#TEST1_9) | [JVET-AB0164](https://jvet-experts.org/doc_end_user/current_document.php?id=12091) | [S. Eadie](mailto:seadie@qti.qualcomm.com) | Y. Li (Bytedance) | **Training** and inference |
| [Test 1.10](#TEST1_10) | [JVET-AB0136](https://jvet-experts.org/doc_end_user/current_document.php?id=12063) | [J. N. Shingala](mailto:jay.shingala@ittiam.com), [T. Shao](mailto:Tong.Shao@dolby.com) | Y. Li (Bytedance) | Only inference |
| NN-based post-filters | | | | |
| [Test 1.11](#TEST1_11) | [JVET-AB0048](https://jvet-experts.org/doc_end_user/current_document.php?id=11963) | [M. Santamaria](mailto:maria.santamaria_gomez@nokia.com) | J. Ström (Ericsson)  J. Li (Bytedance) | **Training** and inference |
| Super-resolution NN-based post-filters | | | | |
| [Test 2.1](#TEST2_1) | [JVET-AB0102](https://jvet-experts.org/doc_end_user/current_document.php?id=12029) | [J. Nam](mailto:junghak.nam@lge.com) | K. Andersson (Ericsson) | Only inference |
| [Test 2.2](#TEST2_2) | [JVET-AB0098](https://jvet-experts.org/doc_end_user/current_document.php?id=12025) | [R. Chang](mailto:renjiechang@tencent.com) | [D. Liu](mailto:du.liu@ericsson.com)  (Ericsson) | **Training** and inference |
| [Test 2.3](#TEST2_3) | [JVET-AB0076](https://jvet-experts.org/doc_end_user/current_document.php?id=11991) | [Q. Han](mailto:hanqihui2013@163.com) | D. Liu  (Ericsson) | Only inference |
| [Test 2.4](#TEST2_4) | [JVET-AB0093](https://jvet-experts.org/doc_end_user/current_document.php?id=12020) | [S. Huang](mailto:shimin_huang2022@163.com) | D. Liu  (Ericsson) | Only inference |
| NN-Intra coding | | | | |
| [Test 3.1](#TEST3_1) | [JVET-AB0125](https://jvet-experts.org/doc_end_user/current_document.php?id=12052) | [F. Galpin](mailto:franck.galpin@interdigital.com) | [Takeshi Chujoh (Sharp)](mailto:chujoh.takeshi@sharp.co.jp) | Only inference |
| [Test 3.2](#TEST3_2) | [JVET-AB0149](https://jvet-experts.org/doc_end_user/current_document.php?id=12076) | [T.Dumas](mailto:thierry.dumas@interdigital.com) | J. Ström (Ericsson) | **Training** and inference |

# Timeline

**T1 - 2 weeks after JVET-AB meeting (11-Nov-2022):** To revise EE description. Changes should be discussed and agreed on JVET reflector.

**T2 – 3 weeks after JVET-AB meeting (18-Nov -2022)**: Initial software release (which includes training scripts that matches what was proposed to the meeting. Anchors (including technologies in NNVC-3.0) are available. For proposals supposed to have training verified training scripts and SW expected not to be modified after T2.

**T3 – 5 weeks after JVET-AB meeting (02-Dec -2022)**: Sufficient explanation about training scripts has been provided to the cross-checker and training verification starts (tests with training stage cross-check only).

**T4 - ~4 weeks before T6 (13-Dec-2022):** Software is frozen (and may include improvements), technology description is ready, test results and complexity analysis are uploaded together with SW; cross-check of compression performance test starts.

**T5 – 3 days before T6 (06-Jan-2023):** Cross-checkers report status to EE1 coordinators (sending e-mail).

**T6 – 10-Jan-2023:** EE1 summary is uploaded as input contribution

# References

1. [] Common Test Conditions and evaluation procedures for neural network-based video coding technology, JVET-AB2016.

   # Links to SW

   <https://vcgit.hhi.fraunhofer.de/jvet-ab-ee1/VVCSoftware_VTM/-/tree/EE1-1.1>~1.11

   <https://vcgit.hhi.fraunhofer.de/jvet-ab-ee1/VVCSoftware_VTM/-/tree/EE1-2.1>~2.4

   <https://vcgit.hhi.fraunhofer.de/jvet-ab-ee1/VVCSoftware_VTM/-/tree/EE1-3.1>~3.2 [↑](#endnote-ref-2)