* **Dataset**

1. The NIC\_Dataset can be accessed at <https://www.bitahub.com/dataset>, which is opened as a public dataset.
2. **Test set.** 1) The test set in the folder of /test contains 24\*4=96 images with 4 different resolutions (ClassA\_6K, ClassB\_4K, ClassC\_2K, ClassD\_Kodak). 2) The test set in the folder of /test\_crop is used to boost the test speed, which is cropped to a resolution of no more than 720p from the original test images at the center position. Note that the final submitted techniques should be tested using the original test images in the folder of /test.
3. **Training set and validation set.** The training set in the folder of /train contains 607,714 256x256 patches, which are cropped from the 1,600 original images and the 2x and 4x down-sampled versions using bicubic interpolation. Similarly, the validation set in the folder of /validation contains 169,798 256x256 patches, which are cropped from the 293 original images and the 2x and 4x down-sampled versions using bicubic interpolation.

The original images of the dataset can be accessed at:

<https://pan.baidu.com/s/1dPTg9JRh4PS748zxdCUUtA> extract code: p76h

* **Train and test protocol：**

1. **Test:**

* **Anchors.** BPG（0.9.8），VVC (VTM-8.0，Intra coding config)
* **Target bitrates.** There are six bitrates：0.06，0.12，0.25，0.5，1.0，1.5bpp.

1. **Fixed-rate model.** Six different models shall be trained using six different λ values, which can reach the six target bitrates for the average bitrate of test set. The maximal allowable bitrate deviation from the target bitrate should be not exceed 5%. Please refer to the algorithm document of NIC-0.1 to get the initial λ values.
2. **Variable-rate model.** Each test image shall be compressed to the 6 target bitrates. The maximal allowable bitrate deviation from the target bitrate should be not exceed 5%.

* **Objective quality assessment.** Four quality metrics will be used: RGB PSNR, YUV 4:2:0 PSNR, YUV 4:2:0 MS-SSIM, and YUV 4:2:0 MS-SSIM (dB). The color space conversion between RGB and YUV should be based on BT.709. The six bitrates will be divided into two ranges: lower range (0.06-0.5bpp) and higher range (025-1.5bpp). For each range, BD-rate will be calculated with respect to each quality metric over each anchor.

YUV PSNR = (6\*Y PSNR+U PSNR + V PSNR)/8.

YUV MS-SSIM = (6\*Y MS-SSIM+U MS-SSIM + V MS-SSIM)/8.

YUV MS-SSIM (dB) = -10\*log10(1- YUV MS-SSIM)

* **Subjective quality assessment**. Same as the CfE。
* **Encoding and decoding time.** The encoding time is calculated as the total time on CPU and GPU from reading the input image to writing out the bitstream. The decoding time is calculated as the total time on CPU and GPU from reading the bitstream to writing out the reconstructed image. The information about the processor platforms should be specified in the proposal, including CPU/GPU brand, number and clock rate of cores, memory usage, and ML framework (eg. Caffe, Tensorflow, Pytorch) and its version information.

1. **Train:**

* **Fixed-rate model.** The loss function (using RGB MSE and RGB MS-SSIM as distortion term), λ values, optimizer, learning rate, batch size, patch size and number of iteration should be same as the default setting of NIC-0.1 unless there are special needs.
* **Variable-rate model**. The hyper parameters for training except for the λ values or quantization parameters should be same as those for fixed-rate model.