ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG16)

Coding of Still Pictures

JBIG
Joint Bi-level Image Experts Group

JPEG
Joint Photographic Experts Group

TITLE: Submission Instructions for the JPEG AI Call for Proposals

SOURCE: Requirements

STATUS: Final

REQUESTED ACTION: Distribute

DISTRIBUTION: Public

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch
Submission Instructions for the JPEG AI Call for Proposals

1. INTRODUCTION

Each team has a folder for uploading their materials (software, images, instructions, etc.) for each task in CfP response. Folder name is TEAMID which is the remote working folder (virtual root) after the ftp login.

Please avoid mentioning names of submitter and organization in file names, comments in a code or/and decoder log; use TEAMID if needed.

For each image processing and computer vision task of your submission there is a dedicated folder inside TEAMID starting with a “_”. Submission structure for these tasks is basically the same as for the standard reconstruction task, namely bit, decoder, rec, rec_by_proponent, rec_by_crosscheck, script files, etc, that appear on the root folder.

The next Sections describe the folder submission structure and some instructions for standard reconstruction, image processing and computer vision tasks.

2. FOLDER SUBMISSION STRUCTURE

The CfP response submission should be organized into folders. The following Figure and Table illustrates the submission folders and a short description of each one.
3. STANDARD RECONSTRUCTION INSTRUCTIONS

JPEG AI naming convention: Mandatory for bitstreams and reconstructed images and should be honored.

- For bit-streams:
  \(<TEAMID>_<IMGID>_TE_<BR>.bits\)

- For bit-reconstructed images
  \(<TEAMID>_<IMGID>_TE_<RES>_<ORIGINAL_BIT_DEPTH>bit_sRGB_<BR>.png\)

RD performance assessment: Please upload (May, 2nd) results reporting template filled with your test data together with bit-streams (“bit” folder) and decoded images (“rec_by_proponent”
The currently available results reporting template contains anchor data for the CfE test set. Anchor data will be up-dated shortly after the “hidden” CfP test set is released.

The JPEG AI Quality Assessment Framework must be used to compute all quality metrics. File “summary.txt” (or “summary.csv”) will appear after metrics computation is done. The content of this file should be inserted into place holder (Fig. 1) of “test” page of reporting template. Place holders for data which proponent should enter manually in “test” and “summary” pages are highlighted in blue in the results reporting template.

![Fig. 1. Place holder for results on “test” page of reporting template.](image)

After the place holder is filled with your test data, all cells on the “summary” page (Fig. 2) are automatically filled, except columns for “S”, “T” (highlighted in blue), which should be filled manually by the proponent. Column “Model” is for the largest number of NN parameters for decoding one bit-stream. Colum “Models” for total number of NN parameters needed for decoding all streams (“all rates”).

Please note, you can select any test as reference, but changing cell “A2” (HEVC is set as reference only for example).

![Fig. 2. “Summary” page of reporting template.](image)

Complexity assessment: kMAC/pxl statistics are mandatory and should be computed during the decoding process and also submitted on 2nd of May. This information should be collected for each rate point and for each test image. Example for kMAC/pxl computation can be found here. This information should be inserted in the test page of the reporting template (column “kMAC/pxl” in place holder of Fig. 1).

Reporting decoding run time for both CPU and GPU is mandatory and the results should be inserted in the test page of the reporting template (columns “DecGPU” and “DecCPU” in place holder, Fig. 1).
Also, please run at least one of the anchors on the same machine as the proposal and place encoding/decoding run-time for this anchor in reporting template (this is needed to have reliable time measurement).

In case the results of decoding on CPU and GPU are not identical, please make a comment to indicate which decoder configuration should be used for the subjective evaluation procedure (“viewing”).

4. **COMPRESSED DOMAIN SUPER RESOLUTION INSTRUCTIONS**

**JPEG AI naming convention:** Mandatory for bitstreams and reconstructed images and should be honored.

- For bit-streams (in bit folder)
  `<TEAMID>_<IMGID>_TE_<BR>.bits`
- For bit-reconstructed images (in rec folder)
  `<TEAMID>_<IMGID>_TE_<RES>_ORIGIN_BIT_DEPTH>bit_sRGB_<BR>.png`

Here is `RES` resolution of up-sampled to full size ground truth image (not resolution of image encoded); `BR` takes values 006, 012, 025, 050, 075.

Same streams must be decodable by decoder in standard reconstruction task submitted by same team, to produce reconstructed image for further decoded anchor computation.

5. **COMPRESSED DOMAIN DENOISING INSTRUCTIONS**

**JPEG AI naming convention:** Mandatory for bitstreams and reconstructed images and should be honored.

- For bit-streams (in bit folder)
  `<TEAMID>_<IMGID>_<NOISE_LEVEL>_TE_<BR>.bits`
- For bit-reconstructed images (in rec folder)
  `<TEAMID>_<IMGID>_<NOISE_LEVEL>_TE_<RES>_ORIGIN_BIT_DEPTH>bit_sRGB_<BR>.png`

Here `NOISE_LEVEL` indicates noise level of encoded image; `BR` takes values 012, 025, 050, 075.

Same streams must be decodable by decoder in standard reconstruction task submitted by same team, to produce reconstructed image for further decoded anchor computation.

6. **COMPRESSED DOMAIN IMAGE CLASSIFICATION INSTRUCTIONS**
**JPEG AI naming convention:** Mandatory for bitstreams and reconstructed images and should be honored.

- for bit-streams (in bit folder)
  
  `<TEAMID>_<IMGID>_TE_<BR>.bits`

Same streams must be decodable by decoder in standard reconstruction task submitted by same team, in order to produce reconstructed images for further decoded anchor computation.

Here BR takes values 012, 025, 050, 075. Top-1 and Top-5 accuracy for each bit-rate should be computed by the sub-task decoder. Examples for Top-1 and Top-5 accuracy computation can be found here.

In this task IMGID is image name in of **ILSVRC 2012**.

For example, names of input images are

```
DATASET_DIRECTORY
n01440764
  ILSVRC2012_val_00000293.JPEG
  ILSVRC2012_val_00002138.JPEG
  ...

n01443537
  ILSVRC2012_val_00000236.JPEG
  ILSVRC2012_val_00000262.JPEG
  ...

...```

Corresponding names for bitstreams are

```
_CLASSIFICATION/bit
n01440764
  <TEAMID>_00000293_TE_<BR>.bits
  <TEAMID>_00002138_TE_<BR>.bits
  ...

n01443537
  <TEAMID>_00000236_TE_<BR>.bits
  <TEAMID>_00000262_TE_<BR>.bits```

In total 50000 images must be processed. Some example of scripts can be found on JPEG AI gitlab: [https://gitlab.com/wg1/jpeg-ai/jpeg-ai-anchors/-/tree/main/Classification](https://gitlab.com/wg1/jpeg-ai/jpeg-ai-anchors/-/tree/main/Classification). The script usage is as follows

```
python -m Classification.process --Classification.data_dir /path/to/dataset --output /path/to/output_dir```

All the proponents are supposed to

1) Produce decoded anchor:
a. Run standard reconstruction decoder (identical to the CfP submission in image reconstruction category)
b. Run ResNet50 using reconstructed images as input for all mandatory rate points 0.12, 0.25, 0.5, 0.75 bpp
c. Compute top-1 and top-5 accuracy

2) Conduct classification in latent space:
   a. Run decoder for object classification task for all mandatory rate points 0.12, 0.25, 0.5, 0.75 bpp
   b. Compute top-1 and top-5 accuracy

Bit-streams for 1) and 2) are the same. The results for decoded anchor and proposed latent space JPEG AI compressed domain classification task results reporting template.xlsx (insert your numbers into cells highlighted in blue).

7. PROPOSAL DESCRIPTION

All proponents must submit a proposal description by 18th July 2022. This proposal description should be in the form of a JPEG input contribution. In case you need assistance to perform this submission, please contact the JPEG AI ad-hoc group chairs.