

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.

Filename
..
_Classification
_Denoising
_SuperResolution
bit
decoder
encoder
rec
rec_by_crosscheck
rec_by_proponent
training
decode.sh
decode_list.sh
encode.sh
encode_list.sh
README.md
requirements.txt
requirements_cpu.txt
train_models.sh

File/directory	Content	Comments
_Classification	package for object classification task	Fill only if you submit response in this category
_Denoising	package for de-noising task	Fill only if you submit response in this category
_SuperResolution	package for super resolution task	Fill only if you submit response in this category
bit	for writing hidden test set bit streams	Write access to this folder will be closed after May, 4 (CfP 2 nd phase).
decoder	contains <u>ALL</u> materials (SW and models) needed for decoding	Write access to this folder will be closed after March, 14. (CfP 1 st phase). <u>NO</u> modifications are possible after.
encoder	contains <u>ALL</u> materials (SW and models) needed for encoding	Write access to this folder will be closed after July, 18 (CfP 3 rd phase).
training	contains <u>ALL</u> materials (SW and models) needed for decoding	Write access to this folder will be closed after July, 18 (CfP 3 rd phase).
rec	Quality assessment framework takes reconstructed images from here	Filled during decoding
rec_by_crosscheck	Folder for reconstructed images	Empty, will be filled by cross-checker
rec_by_proponent	Folder for reconstructed images	Submitted by proponent together with bit-streams
decode.sh	Example of command line for decoder. CPU decoding must supported, GPU recommended to be supported	decode.sh reads streams from "bit" folder, produces reconstructed images and write them to "rec" folder. Produces txt and json files with decoder run-time and kMAC/pxl.
decode_list.sh	Calls decoder.sh for list images in hidden test set	./decode_list.sh cpu → for CPU decoding ./decode_list.sh gpu → for GPU decoding
encode.sh	Example of command line for encoder. CPU encoding must supported, GPU recommended to be supported	encode.sh reads test set images, produces bit-streams and writes them to "bit" folder. Produces txt and json files with encoder run-time.
encode_list.sh	Calls encoder.sh for list images in hidden test set	./encode_list.sh cpu → for CPU decoding ./encode_list.sh gpu → for GPU decoding
train_models.sh	Training scripts	Scripts for training: from reading and processing training data until model is produced. The training scripts should produce a model that is made available in the "decoder" folder.
README.md	File with instructions for all tasks	

3. STANDARD RECONSTRUCTION INSTRUCTIONS

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

- For bit-streams:

<TEAMID>_<IMGID>_TE_
.bits

- For bit-reconstructed images

<TEAMID>_<IMGID>_TE_<RES>_<ORIGINAL BIT DEPTH>bit_sRGB_
.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")

folder). 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.

EU	EV	EW	EX	EY	EZ	FA	FB	FC	FD	FE	FF	FG	FH	FI	FJ	FK	FL
TEAMID																	
name	bpp	mssim Torc	mssim lqa	psnrY	psnrU	psnrV	vif	fsim	nlpd	iw-ssim	vmaf	psnrHVS	kMAC/pxl	DecGPU	DecCPU	EncGPU	EncCPU
TEAMID_																	

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. Column “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.

Reference:																				
HEVC																				
	5 points BD-rate (0.06, 0.12, 0.25, 0.5, 0.75)																			
	BD rate vs HEVC									Dec. complexity						Enc. complexity				
Test	msssim	Torc	vif	fsim	nlpd	lw-ssim	vmaf	psnrHVS	taxBitDl	MAX kMAC/pxl	AVG kMAC/pxl	Time GPU	Time CPU	Model	ModelS	Time GPU	Time CPU			
JPEG2000	59.6%	73.9%	68.1%	46.7%	65.2%	63.3%	30.7%	69.5%	2%	0	0	0.1	0.1	0	0	0.0	0.0			
HEVC	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	11%	0	0	1.0	1.0	0	0	1.0	1.0			
VVC	-11.4%	-9.6%	-11.4%	-17.7%	-10.0%	-11.1%	-11.9%	-8.3%	14%	0	0	1.4	1.4	0	0	3.6	3.6			
TEAMID	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	0	0	0.0	0.0	0	0	0.0	0.0			

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_
.bits

- for bit-reconstructed images (in rec folder)

<TEAMID>_<IMGID>_TE_<RES>_<ORIGINAL BIT DEPTH>bit_sRGB_
.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 imaged 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_
.bits

- For bit-reconstructed images (in rec folder)

<TEAMID>_<IMGID>_<NOISE LEVEL>_TE_<RES>_<ORIGINAL BIT DEPTH>bit_sRGB_
.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 imaged 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_
.bits`

Same streams must be decodable by decoder in standard reconstruction task submitted by same team, in order to produce reconstructed image 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>. The script usage is as follows

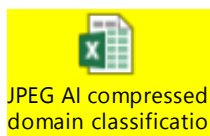
```
python -m Classification.process --Classification.data_dir /path/to/dataset --
output /path/to/output_dir
```

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.