EE838A

Special Topic on Image Engineering: Advanced Image Restoration and Quality Enhancement

Homework Assignment 4

Implementation and Verification of deep learning based video frame interpolation networks

The fourth homework assignment is to implement a given deep neural network for video frame interpolation and to train it properly. After training, the performance verification is required for test images by providing the performance measures in terms of PSNR, SSIM and MS-SSIM (Multi-scale SSIM). The target implementation of deep frame interpolation is the following paper which has been covered in the class:

> Z. Niklaus, L. Mai, and F. Liu, "*Video Frame Interpolation via Adaptive Separable Convolution*," ICCV, Venice, Italy, Oct. 2017.

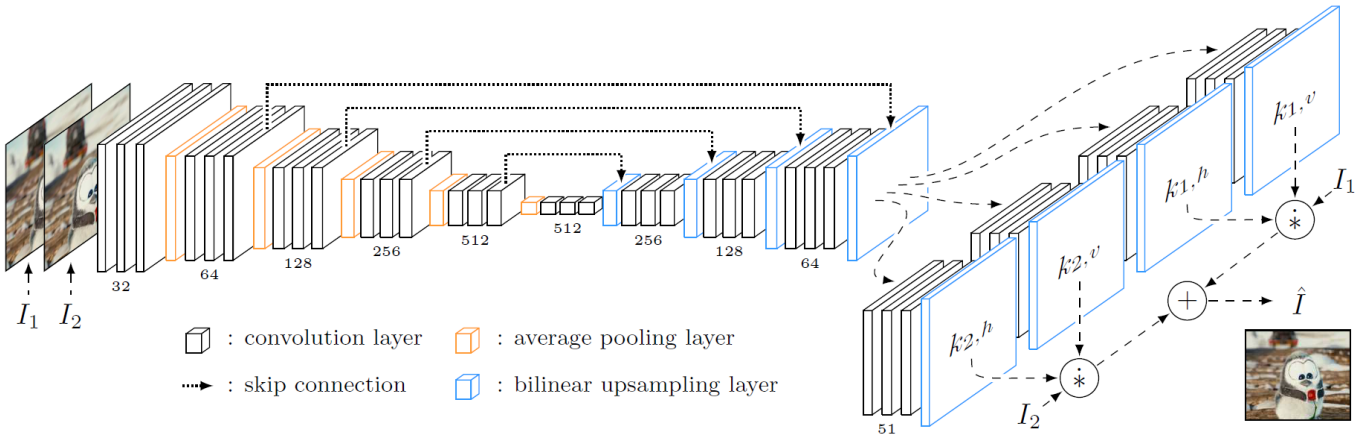
The data set for experiments is provided via the below Dropbox link. The evaluation of the submitted trained networks for Homework Assignment 4 will be performed based on the correct implementation and performance of the given network.

Dataset

- Training and validation samples: Youtube dataset (5 videos with 720p 60fps)
- You should clip the scene changes in the video to use them as training data.
- Download links:
  - Training/Validation data: https://www.dropbox.com/s/wfci581y8tgyagr/60fps.zip?dl=0

Implementation

- Adam Optimizer
- L1 loss (Optional - Perceptual loss : relu4_4 layer of the VGG-19 network)
- Training patch size : 128×128×3 RGB images, batch size : any numbers
- You can use several data augmentation techniques (like randomly flipping horizontally and vertically, rotating by 90 degrees)
- All of the convolution filter size is 3×3 with ReLU.
- Other parameters for training is freely selectable.
- The details of the training conditions can be found in the paper (N. Simon et al. “Video Frame Interpolation via Adaptive Separable Convolution,” ICCV, 2017.)
- The following multi-scale network architecture for the video frame interpolation



$$\hat{I}(x,y) = k_{1,h} * \left[k_{1,v} * I_1(x,y)\right] + k_{2,h} * \left[k_{2,v} * I_2(x,y)\right]$$

As shown in the above equation, the network is trained to learn the four 1-D kernels, $k_{1,v}$, $k_{1,h}$, $k_{2,v}$, and $k_{2,h}$.

The following deliverables must be submitted:

- Both training and test codes (Tensorflow or Pytorch)
- Readme.txt (A simple description of how to execute your codes)
- Report (including the interpolated frames produced by your test code for the three validation frames, analysis of your results and a simple code description for each component of the neural network)

Submission

- Due date: 2018-12-01 (Saturday) 23:59
- Submission should go to the class TAs at: shki@kaist.ac.kr and pys5309@kaist.ac.kr
- Submission format
  - Your report must include your name, student ID, phone number and e-mail
  - Your report must be in ZIP format with following directories:
    - ✓ `source` where readme.txt, training code and test code must reside
    - ✓ `report` where your report is put
  - The file name of your submission should be "HW4_studentID_YourName.zip".

Evaluation of homework:

- The inference codes that the students submitted will be tested for unknown video sequences and their performances will be measured for evaluation of the homework.