IE643: Deep Learning - Theory and Practice **Jul-Dec 2023**

Assignment 3: Due On 31st October 2023 (11:59 PM IST)

1 Instructions

Answer all questions. Write your answers clearly. For every exercise problem, create a separate python notebook and make sure that all your answers are present in that notebook. Name your python notebook files as `IE643_YOURROLLNO_assignment3_Ex1.ipynb`, `IE643_YOURROLLNO_assignment3_Ex2.ipynb` and `IE643_YOURROLLNO_assignment3_Ex3.ipynb`.

Construct a folder named `IE643_YOURROLLNO_assignment3` which contains all the notebook files. Upload on moodle, a single zip file named `IE643_YOURROLLNO_assignment3.zip`, containing the folder with all the notebook files.

IMPORTANT NOTE: Make sure that your code can be run in Google colab as it is. Include the installation code for all relevant packages used in your code. No further installations should need to be performed by the TAs during evaluation.

You can score a maximum of 50 marks in this assignment.

There will be no extensions to the submission deadline.

2 Data set preparation

Download the data set `Assignment3_dataset.zip` from the link shared in moodle. **Please use IITB Google SSO login to download the data**.

Unzip the data into your computer (`e.g.` you can use `unzip Assignment3_dataset.zip` if you use linux).

The data set contains scenes captured in low light situations (e.g. scenes captured at nights, early mornings or sunsets), containing objects belonging to 3 different categories namely: boat, bus and car. These images are present in 3 different folders with appropriate names. Note that there are roughly 110 images in each category. Further note that the images are of different resolutions (height, width). For your answers, use a consistent size of images (e.g. 32×32, 64×64 etc.) based on the memory requirements, network choices and other considerations.

You are free to use data augmentations (e.g. random crops, flips along vertical axis, minor deformations by inducing structured noise, small-angle roations) whenever required.

You are also free to use deep convolutional networks of your choice. However you would need to justify the choices made.

3 Questions

1. **[Use only Python]** Choose a category C of your choice and construct a data set D containing 90% images from the category C of the data. Construct a validation set V using the remaining 10% images.

   The first question is about constructing a Deep Convolutional GAN to generate images similar to those in category C. Please note that you need to choose only a single category (either boat or bus or car) for this exercise. Answers generating all three categories of objects will not be considered.

   (a) Construct a Deep convolutional GAN (DC-GAN). You can use the demo code posted in moodle to construct your DC-GAN. You are free to use other CNN architectures for your GAN based on the requirements of the data set. If you choose a different architecure, you must clearly justify your choice for the same.

   (b) Train your DC-GAN on the data D.

   (c) Choose the best parameters for DC-GAN (e.g. number of training iterations, learning rate, batch sizes to train the discriminator and generator, etc.) using the validation set V.

   (d) If needed, use heuristics from the paper [1] to improve training. You can use early stopping criteria and learning rate scheduling based on your choice.

   (e) Clearly describe in your python notebook about the training heuristics used in your code and explain why they were useful.

   (f) Prepare a plot of the training objective function value against the iterations.

   (g) Prepare a plot to depict the discriminator objective function value and the generator objective function value and check if they converge.

   (h) After training, display 10 images (in a 10×1 grid) generated using the generator of your DC-GAN. For each of the 10 images, display 5 original images from the training data D closer to the generated image in a 10×5 grid. Comment on the quality of the generated images. Explain the similarity metric you used in the code and justify your choice.

   (i) IMPORTANT: Make sure that the images you have generated using your DC-GAN are meaningful real-looking images. Answers having no meaningful image generation will not be considered for evaluation.

2. For this question, choose a different object category K, different from C chosen in the previous question. Construct a data set D_1 containing 90% images from the category K of the data. Construct a validation set V_1 using the remaining 10% images. This question is about constructing a VAE to generate images similar to those in category K. Please note that you need to choose only a single object category for this exercise. Answers generating all three categories of objects will not be considered.

   (a) Construct a suitable VAE architecture using convolutions and related operations (e.g. upsampling, pooling, batch normalization etc). You can use the demo code posted in moodle to construct your VAE. You are free to use other CNN architectures for your VAE based on the requirements of the data set. If you choose a different architecure, you must clearly justify your choice for the same.

   (b) Train the VAE on the data D_1.

   (c) Choose the best parameters for VAE (e.g. number of training iterations, learning rate, weightage for KL term, etc.) using the validation set V_1. You can use early stopping criteria and learning rate scheduling based on your choice.

[1] `https://arxiv.org/pdf/1511.06434.pdf`

(d) Clearly describe in your python notebook about the training heuristics used in your code and explain why they were useful.

(e) Clearly describe the objective used for VAE and your training and inference procedures.

(f) Prepare a plot of the training objective value against the iterations.

(g) After training, display 10 images (in a 10×1 grid) generated using the VAE. For each of the 10 images, display 5 original images from the training data D_1 closer to the generated image in a 10×5 grid. Comment on the quality of the generated images. Explain the similarity metric you used in the code and justify your choice.

(h) IMPORTANT: Make sure that your VAE is useful for generating real-looking images. Answers having no meaningful image generation will not be considered for evaluation.

3. Use the DC-GAN you developed in Question (1) to train on category K and generate 10 images. Compare the quality of images generated using DC-GAN and VAE for category K. Similarly, use VAE in question (2) to train on category C and generate 10 images. Compare the quality of images generated using VAE and DC-GAN for category C. Explain your observations and the possible reasons for your observations.

END OF ASSIGNMENT QUESTIONS