CENG 796 - Peer-review form

Reviewed project ID: Group 5

Reviewed project's title (title of the paper): Dual Contradistinctive Generative Autoencoder

Reviewer name(s): Sezai Artun Ozyegin – Merve Tapli

Instructions:
- Answer = *Yes*, *No* or *Partial*.
- You may expand sections as necessary.
- For most questions, you do not need to add comments, unless the instructions tell you otherwise.
- "Notebook" refers to "Jupyter Notebook" file that is expected to be named as main.ipynb

Question	**Answer**	**Comments**
Contains a jupyter notebook file	Yes	
Notebook is located at <project_root>/main.ipynb	Yes	
Notebook's first section contains paper information (paper title, paper authors, and project group members' name & contact information) Some good examples: see group03, group10, group11 (and a couple of other groups).	Yes	Additionally, there was a nice paper summary and a section about loss functions.
Notebook contains a section for hyper-parameters of the model.	Yes	
Notebook contains a section for training & saving the model.	Yes	
Notebook contains a section (or a few sections) for loading a pre-trained model & computing qualitative samples/outputs.	Yes	
Notebook contains reproduced plots and/or tables, as declared.	Yes	
Notebook contains pre-computed outputs.	Yes	

Data is included and/or a proper download script is provided.	Yes	The dataset is downloaded using torchvision and the saved model is included in the submission. Even there is a script for downloading the saved model from Google Drive if it is not available.
Notebook contains a section describing the difficulties encountered.	Yes	In the challenges section, it is said that they need to read the MoCo paper. However, we think that the whole formulation is given in the original paper and could not understand the necessity of it.
The paper has achieved its goals and/or explained what is missing.	No/Yes	The sampled examples qualitatively seem nice. However, the FID score did not meet the goal. The reconstructed examples are not similar to the original ones. It was not included in their goals; however, as discussed in the challenges section, it may indicate there is a problem.
The notebook contains a section that reproduces the figure(s) and table(s) declared in the goals.	Yes	
The notebook also reports the original values of the targeted quantitative results, for comparison.	Yes	
MIT License is included.	Yes	
As the reviewer(s), you have read the paper & understood it.	Yes	
Implementation of the model seems correct.	Yes	The model and loss functions seem correct. The all three architectures match with the ones given in Figure 15. There were some minor issues: • ResConvBlock always skips the pre-activation layer. • In the Discriminator module, the cont_conv layer with kernel size 1x1 has padding, which seems redundant. • We could not find the ‘local patch’ implementation. (Section 4.2) There were also some parts that were ambiguous (or we may have missed them) in the papers (DC-VAE and AutoGAN), but some decisions had been taken during the reimplementation: • In ResConvBlocks, a 1x1 conv layer is used over every identity connection. In original ResNet implementation, 1x1 conv layers are used only when downsampling is applied. It may be different for AutoGAN. • For the Encoder and Discriminator, we could not find any information about where to downsample. In the reimplementation, the downsampling is applied in the first 2 ResConvBlocks. • After some of the skip layers in Decoder, 1x1 conv layer was applied. We also could not find that in the paper. • In the last block of Decoder (to_rgb), BatchNorm is used. Again, we could not find any info about it. • In the contrastive head, ReLU is applied after the cont_lin linear layer. However, this is not the case for the other

		branch. Also, the paper does not mention how to merge them. In the reimplementation, the branches are merged with sum, which seems appropriate. • In the paper the output channel size of the 1x1 conv applied for 'deep supervision' is not mentioned. The reimplementation uses it as 1, which may or may not be enough.
Notebook looks professional (in terms of notation, readability, etc.)	Yes	We don't have any suggestions.
Source code looks professional (in terms of coding style, comments, etc.)	Yes	

Additional comments:

- In the notebook it is said that the batch size is not mentioned in the paper. However, batch size is given as 128 in section 5.1 of the paper. The reimplementation uses 64. It may be due to memory constraints.
- The paper experiments on the number of negative samples to be used in the contrastive loss in Figure 7. They tried values in between 128 and 8096 and decided to use 8096. In the Figure 7 the results seem close, but the reimplementation chooses this size to be the same as the batch size (64).
- The paper calculates the FID over 50000 samples. In the reimplementation it is 10000. Due to the batch normalization, it may affect the results.
- To calculate the FID, the reimplementation saves the images with cv2.imwrite, which uses 'BGR' as its default convention. However, the generated images are in 'RGB' space. Therefore, the images are saved with the wrong channel order and Inception network takes these images. We are not sure how real CIFAR10 images' mu and sigma are calculated.