

Description

This homework is dedicated to the pix2pix model[1]. The model was created for the DL course I had few years ago.

The project contains:

- Generator
- Training only the generator with L1 loss
- Discriminator
- Training both generator and discriminator together
- Testing the model on the 2 datasets: facades (the authors of the initial paper created this dataset) and the dataset with Funko Pops - this dataset was created solely by me for this homework (I can provide the dataset, the preprocessing and raw data I collected).
- Experiments with augmentations, but they did not make to the final notebook with GAN, but you can see some code and examples in the report (in plot_transforms notebook in which I checked how different transforms change my pictures to find the hyperparameters which could be potentially used)

In the funkopop_final_GAN.ipynb notebook you can find the models and functions + experiments on my dataset. Pictures for experiments with facades are only in the report (it is in Russian, but there are a lot of pictures)

Also here are some reports from weights and biases (there are not many runs shown, I created other projects while experimenting, these are just some of them) <https://wandb.ai/kogana00/facades%20GAN?workspace=user-kogana00> and <https://wandb.ai/kogana00/funkopops%20GAN?workspace=user-kogana00>

Overlap with course - pix2pix is a DL model, it contains such blocks as Conv2d, activation functions (LeakyReLU), normalization, Dropout. I used optimizers, loss function, validation set, etc (a common DL training pipeline), experimented with augmentation

Conclusion

Already done:

- Code for models, training and testing is written
- Experiments with hyperparameters, augmentations are conducted
- Own dataset is collected
- The final models are tested

I think I have already covered more than required 50% of work.

What I plan to do:

- Refactor code: It was all written in one notebook with enormous outputs (we were forbidden from using any libraries for tracking progress, so I just printed some info), so it is hard to read now.
- Rewrite the old report in English
- Train the models again: I am not sure if I can find old files with models
- (Optional) Try another dataset: The authors of the initial paper[1] provided several datasets. However, training the model takes a lot of time even on GPU and I am not sure if I will have enough time (Colab gives short access to GPUs)

References

- [1] Phillip Isola and Jun-Yan Zhu and Tinghui Zhou and Alexei A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, 2018