BLG453E
Homework-4

03.02.2022

Res. Asst. Yusuf Huseyin Sahin
sahinyu@itu.edu.tr

- You should write all your code in Python language.
- Cheating is highly discouraged. If you are planning to use different libraries or functions, please ask me about it.
- Ninova only stores files under 20 MB. If you could not upload your results, you can share them with me via Dropbox, or send me private YouTube video links for each part's results.

1 - Part 1: (80 pts) Motion

I'm standing on the edge of some crazy cliff. What I have to do, I have to catch everybody if they start to go over the cliff—I mean if they're running and they don't look where they're going I have to come out from somewhere and catch them. That's all I'd do all day. I'd just be the catcher in the rye and all.

Catcher in the Rye, J. D. Salinger (1951)

Part 1.1 (50 points): Motion Estimation
I uploaded a frame sequence showing the main gate of the university in a top-down view. You can download the sequence using this link[1]. A frame from the sequence is given in Figure 1. As can be seen from the figure, there are three people walking.

Implement Lucas-Kanade algorithm locally to find the OF vector for these three people. Using cv2.arrowedLine function, place one arrow on each person according to the OF vector. Create a video containing arrow placed images. **Do not use any predefined positions for the people. Guess their whereabouts via OF.**

Part 1.2 (40 points): Background Subtraction
For the same image sequence, use one of the still images and do background differencing. Create a video showing these three people as blobs.

[1] https://web.itu.edu.tr/sahinyu/hw4/DJI_0101.zip

Figure 1: A frame from the sequence.

2 - Part 3: (10 pts) I am somebody now!

I'm not a number, I'm a free man!

The Prisoner (1967)

In the second homework, you used StyleGAN generator and obtained face images from random numbers. In this homework, we will do the same thing with face images. Pixel2Style2Pixel Framework [2] creates vectors representing the given face. Using this vectors in the StyleGAN, the same image is obtained. The procedure could be seen in Figure 2.

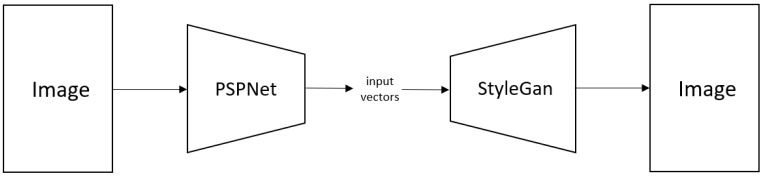

Figure 2: PSPNet and StyleGAN.

3 - BONUS: Dimensionality Reduction (20 pts)

As in the first homework, you will again use the data from ICCV VIPriors Image Classification Challenge (2021)[3]. You will predict the classes of test images on a trained neural network structure as follows:

- For different networks, obtain latent features for each image.
- Since you have many vectors from many networks, you can use dimensionality reduction. You could use other methods to process these vectors too. Heavy image augmentation is also not prohibited.

[2] Richardson, Elad, et al. "Encoding in style: a stylegan encoder for image-to-image translation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

[3] https://competitions.codalab.org/competitions/33214

- For each test image, use k-NN classification on the feature space.

You can find the networks to obtain the features from this link[4]. The kaggle link will be published soon. For the 1^{st} and 2^{nd} place students, I will give two original copies of Deep learning book by Ian Goodfellow, Yoshua Bengio and Aaron Courville.

[4] https://web.itu.edu.tr/sahinyu/hw4/feature_extractors.zip