

COMP4211 PA2 Report

Name: Ruiming Min; SID: 20827430; ITSC: rmin

March 17, 2024

5 Datasets and Data Loaders

5.3 Dataloader Implementation

[Q1]

Number of images in COCO dataset: 3557. Number of images in WikiArt dataset: 7492.

[Q2]

Number of images in PACS train dataset: 1641. Number of images in PACS test dataset: 2723.

6 Style Transfer

6.2 Model

[Q3]

Since after encoding, the image is represented as a 3D tensor and its 2D shape is much smaller than the original image. Therefore, in the decoder, we need the upsampling to re-construct the image to its original size.

[Q4]

For each layer, we have the trainable parameters as follows:

- Conv2d: $(k \times k \times c_{in} + 1) \times c_{out}$
- Pooling and Upsampling: 0

Therefore, the total number of trainable parameters is 2332511.

[Q5]

Since we need an encoder to encode the image and mine the features, but if we only use the encoder, we cannot re-construct the image. Therefore, we need the decoder to re-construct the image.

[Q6]

Since in this model, we do not only need the style match our target style, but also need the content to be preserved. Therefore, we need to use the content loss and style loss to balance the style transfer.

[Q7]

After 55 epochs, total loss is 13021.544, content loss is 8052.103, and style loss is 4969.441.