

可图Kolors-LoRA模型训练报告

姓名：姜雨童
学号：3220103450
报告日期：2024.9.1

Chapter 1: 结果展示

8图故事：bjd娃娃会一直陪你

1.或许今天她坐在秋千上感受微风拂面。



2.或许今天她只是看着窗外发呆。



3.或许今天她带着兔子玩偶漫步街头。



4.或许今天她坐在公园长椅上看蚂蚁搬家。



5.或许今天她是荷叶上超凡脱俗的仙子。



6.或许今天她偶遇了一株比她还大的蘑菇。



7.或许今天她是舞会上耀眼夺目的宝石。

8.在这些不确定里唯一能确定的是，今天的她也会像昨天那样，对你说一声晚安。还有你们的无数个明天，亦是如此。

8图故事：bjd娃娃会一直陪你



Chapter 2: 选用的数据集

因为想要做bjd（Ball—jointed Doll，球型关节人偶）风格的LoRA，所以使用了一组实拍的bjd人偶照片作为数据集（见图2-1，魔搭平台上数据集地址为bjd_test_555·数据集 (modelscope.cn)](https://modelscope.cn/datasets/jyt555/bjd_test_555)）。

考虑到bjd人偶的服饰在视觉呈现上和普通衣物类似，数据集更多侧重于提供bjd人偶在五官/皮肤/面部轮廓走向/不同打光下阴影呈现等层面的训练信息，因此数据集采用的图片为bjd人偶在不同角度/打光下的照片，以及正面特写照。同样地，由于数据集侧重于提供模拟人偶的信息，而非美学方面的信息，我并没有通过代码计算客观美学评分以筛选数据集集中的图片。

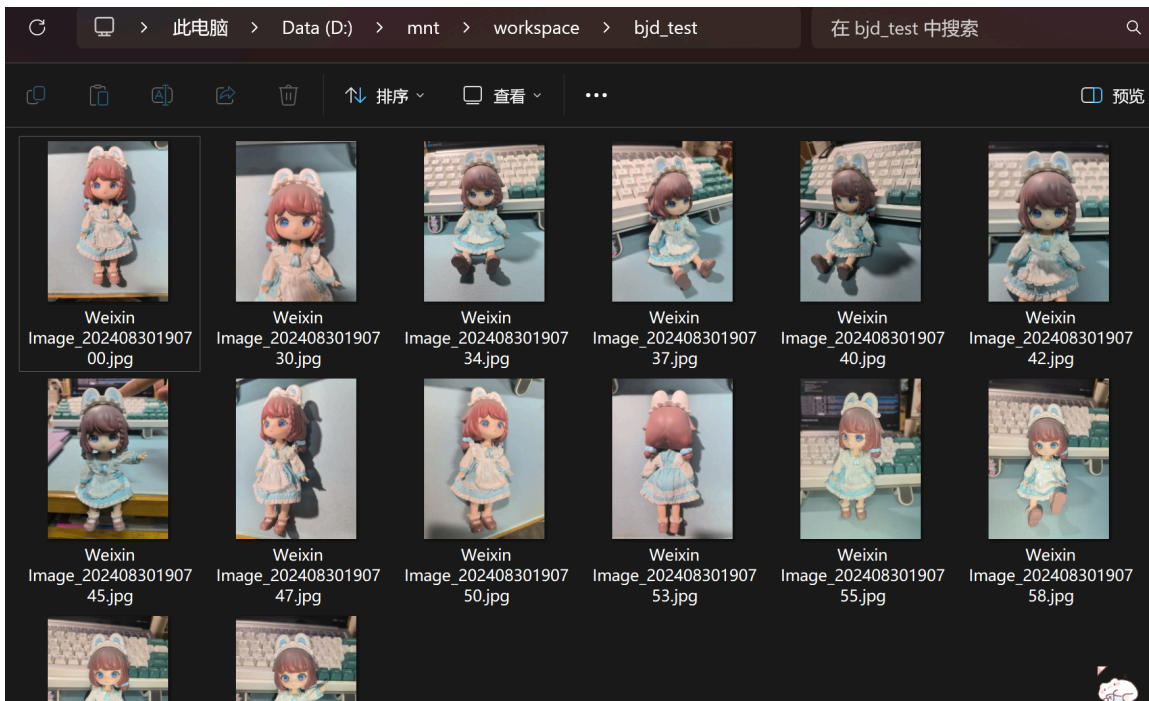


图2-1

Chapter 3: 基础模型

使用的基础模型为：Kolors-SDXL1.0

LoRA训练脚本基本沿用比赛给予的baseline.ipynb，并在此基础上做了微调。

最终bjd-LoRA模型链接：[未命名.docx-可图Kolors训练-云养bjd·模型库 \(modelscope.cn\)](#)

Chapter 4: 微调代码

使用的LoRA训练脚本（报告所在文件夹下555test.ipynb）在原脚本基础上做出如下改动：

1 下载数据集

使用了自行创建的数据集bjd_test_555：

```
1 from modelscope.msdatasets import MsDataset
2 ds = MsDataset.load(
3     'jyt555/bjd_test_555',
4     subset_name='default',
5     split='train',
6     cache_dir="/mnt/workspace/data"
7 )
```

并将触发词修改为“bjd”：

```
1 import json, os
2 from data_juicer.utils.mm_utils import SpecialTokens
3 from tqdm import tqdm
4
5 os.makedirs("./data/lora_dataset/train", exist_ok=True)
6 os.makedirs("./data/data-juicer/input", exist_ok=True)
7 with open("./data/data-juicer/input/metadata.jsonl", "w") as f:
8     for data_id, data in enumerate(tqdm(ds)):
9         image = data["image"].convert("RGB")
10        image.save(f"/mnt/workspace/data/lora_dataset/train/{data_id}.jpg")
11        metadata = {"text": "bjd", "image":
12            [f"/mnt/workspace/data/lora_dataset/train/{data_id}.jpg"]}
13        f.write(json.dumps(metadata))
14        f.write("\n")
```

2 Lora-rank

由于本LoRA侧重于提供bjd风格的图像而非特定人物（8图作品中人物形象统一主要由prompt来实现），因此将Lora-rank修改为较大值20：

```
1 import os
2
3 cmd = """
4 python DiffSynth-Studio/examples/train/kolors/train_kolors_lora.py \
5     --pretrained_unet_path models/kolors/Kolors/unet/diffusion_pytorch_model.safetensors \
6     --pretrained_text_encoder_path models/kolors/Kolors/text_encoder \
7     --pretrained_fp16_vae_path models/sdxl-vae-fp16-
8     fix/diffusion_pytorch_model.safetensors \
```

```

8     --lora_rank 20 \
9     --lora_alpha 4.0 \
10    --dataset_path data/lora_dataset_processed \
11    --output_path ./models \
12    --max_epochs 1 \
13    --center_crop \
14    --use_gradient_checkpointing \
15    --precision "16-mixed" \
16    --modelscope_model_id xxxxx \
17    --modelscope_access_token xxxxx
18    """.strip()
19
20    os.system(cmd)

```

3 prompt

生成8图故事的prompt如下：

```

1  torch.manual_seed(3)
2  image = pipe(
3      prompt="bjd，一个粉色中长发蓝色眼睛的小女孩，穿着白蓝配色的洛丽塔裙子，坐在秋千上",
4      negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
5      cfg_scale=4,
6      num_inference_steps=50, height=1024, width=1024,
7  )
8  image.save("1.jpg")
9
10 torch.manual_seed(4)
11 image = pipe(
12     prompt="bjd，一个粉色中长发蓝色眼睛的小女孩，穿着黄白配色洛丽塔，披散着头发，站在橱窗前",
13     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度，巨大的眼睛",
14     cfg_scale=4,
15     num_inference_steps=50, height=1024, width=1024,
16 )
17 image.save("2.jpg")
18
19 torch.manual_seed(5)
20 image = pipe(
21     prompt="bjd，一个粉色中长发蓝色眼睛的小女孩，穿着白色洛丽塔裙子，头发分两边打着卷，站在大街上，手里抱着白色兔子玩偶",
22     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度，色情擦边",
23     cfg_scale=4,
24     num_inference_steps=50, height=1024, width=1024,
25 )
26 image.save("3.jpg")
27
28 torch.manual_seed(3)
29 image = pipe(
30     prompt="bjd，一个粉色中长发蓝色眼睛的小女孩，双马尾扎得很低，穿着白色为主，浅绿色为辅的洛丽塔裙子，坐在公园长椅上，低头，打着一把白色蕾丝花边太阳伞",
31     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
32     cfg_scale=4,
33     num_inference_steps=50, height=1024, width=1024,
34 )
35 image.save("4.jpg")
36

```

```

37 torch.manual_seed(4)
38 image = pipe(
39     prompt="bjd, 一个粉色中长发蓝色眼睛的小女孩, 穿着纯白洛丽塔裙子, 坐在一片大荷叶上",
40     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
41     cfg_scale=4,
42     num_inference_steps=50, height=1024, width=1024,
43 )
44 image.save("5.jpg")
45
46 torch.manual_seed(1)
47 image = pipe(
48     prompt="bjd, 一个粉色中长发蓝色眼睛的小女孩, 披散头发, 穿着主体红色, 配有白色装饰的洛丽塔裙子, 靠在一朵大蘑菇上",
49     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
50     cfg_scale=4,
51     num_inference_steps=50, height=1024, width=1024,
52 )
53 image.save("6.jpg")
54
55 torch.manual_seed(3)
56 image = pipe(
57     prompt="bjd, 一个粉色中长发蓝色眼睛的小女孩, 穿着淡橘色洛丽塔裙子, 两边打卷, 在舞会上",
58     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
59     cfg_scale=4,
60     num_inference_steps=50, height=1024, width=1024,
61 )
62 image.save("7.jpg")
63
64 torch.manual_seed(3)
65 image = pipe(
66     prompt="bjd, 一个粉色中长发蓝色眼睛的小女孩, 披散着头发, 穿着淡蓝色洛丽塔裙子, 坐在小婴儿床上",
67     negative_prompt="丑陋、变形、嘈杂、模糊、低对比度",
68     cfg_scale=4,
69     num_inference_steps=50, height=1024, width=1024,
70 )
71 image.save("8.jpg")

```

Chapter 5: 分工说明

本报告内所有工作均由本人独立完成, 包括数据集收集、模型训练、prompt编写等。

同组成员邵思捷在本模型基础上进行微调, 并编写了另一套prompt生成她的8图故事。

(同组另两名成员训练了自己的模型, 不参与本模型训练等工作。)