

画像→画像変換の技 Image-to-Image/ ControlNet/IP-Adapter

佐藤 聖

表1 第3章で紹介する技術

技術	詳細
Image-to-Image	既存の画像をベースに、スタイルを変えたり新しい要素を追加したりといった画像編集を行う
ControlNet	生成される画像の特定の要素(ポーズ、線画、深度など)を、外部から提供した画像やデータをもとに制御する。より詳細な画像生成が可能になる
IP-Adapter	既存の画像全体を、新たな画像生成の種として利用する。入力画像のスタイルや特徴を新たな画像に反映させることができる
前章までのText-to-Image	文字列で表現された指示に基づいて画像を生成する

第3章では、既存画像をベースに画像を生成する手法について紹介します。具体的には、表1の3つの技術について実験を交えて解説します。プログラムは本誌ウェブ・ページから入手できます。

<https://www.cqpub.co.jp/interface/download/contents.htm>

実験1 画像スタイルの変更や新要素の追加ができる

プログラム：img2img001～img2img004フォルダ

● 実験すること

Image-to-Image技術に関連したPythonプログラムを紹介します。異常画像を作成する際に、このタスクを利用すると、プロンプトに加えて、元の画像をベースにして、異常箇所を加えた画像を作成できます。この方法は、異常検出モデルの訓練用データを生成する際に非常に役立ちます。

● 実験プログラム

リスト1のように、元画像にプロンプトで指示を加えることで、ユーザの意図に合わせた創造的でリアルな画像を生成します。これにより、少ない労力で写真のようなリアルな画像の作成や編集が可能になり、

リスト1 既存の画像をベースに新たな画像を生成するImage-to-Image

```

1 # ライブラリを読み込み
2 import requests
3 from PIL import Image
4 from io import BytesIO
5 from diffusers import
      StableDiffusionImg2ImgPipeline
6 import torch
7
8 # 画像のダウンロードとリサイズ
9 url = "https://raw.githubusercontent.com/
      CompVis/stable-diffusion/main/assets"
10 response = requests.get(url + '/stable-samples/
      img2img/sketch-mountains-input.jpg')
11 init_image = Image.open(BytesIO(
      response.content)).convert("RGB")
12 init_image = init_image.resize((768, 512))
13
14 # 画像生成パイプラインを設定
15 device = "cuda"
16 pipe = StableDiffusionImg2ImgPipeline.from_
      pretrained("CompVis/stable-diffusion-v1-4",
17 torch_dtype=torch.float16).to(device)
18
19 # プロンプト
20 prompt = "Photo realistic, RAW, fantasy
      landscapes, masterpiece, best quality,
      high quality"
21
22 # パイプラインを実行
23 with torch.autocast(device):
24     image = pipe(prompt, image=init_image,
25                 num_inference_steps=100, strength=0.75,
26                 guidance_scale=7.5).images[0]
27
28 # 入力画像を保存
29 init_image.save("001-1-org.png")
30 # 画像を保存
31 image.save("001-1-output.png")

```

ユーザはイメージを効果的に表現できます。

▶ ライブラリを読み込む

- 2行目…インターネットから画像データを取得します。
- 3行目…PIL (Python Imaging Library) は画像の操作や処理を行うためのライブラリです。
- 4行目…BytesIOは取得したデータをメモリ上で扱うためのものです。
- 5行目…Stable Diffusionの画像生成パイプラインを構築します。
- 6行目…torchはGPUを活用するためのライブラリです。