第3章 インペインティング機能を使って任意の場所に異常を作り込む

レッスン②… 画像の一部を加工する

佐藤 聖

ここでは，もともと正常だった画像の一部を，異常があるように加工して出力します．この加工にはStable Diffusionのインペインティング技術を利用します．インペインティング技術の流れは次の通りです．

1. 修復したい部分や消去したい部分をマスクで覆う
2. マスクされた画像をStable Diffusionの拡散モデルに入力
3. マスクされた部分に新しい画像を生成
4. 生成された画像を元の画像に合成

このプロセスを通じて，本格的に異常画像を作り上げていきます．

ステップ1… 道路のひび割れ画像を作る①

道路の異常には，アスファルトにできた穴やひび割れがあります．ここではひび割れ画像を画像生成AIで作ります．前章で紹介したText-to-Image用のPythonプログラムを元に，インペインティング用のプログラムを作成します．前章のプログラムの流れやコードの構造が理解できていれば，この変更はそれほど難しくないでしょう．基本的な流れを**リスト1**に示します．

リスト1　インペインティング機能を使って画像の一部を加工するPythonプログラム（inpainting-01.py）

```
 1 # ライブラリを読み込み
 2 from diffusers import AutoPipelineForInpainting
 3 from diffusers.utils import load_image
 4 import torch
 5
 6 # 画像生成パイプラインを設定
 7 model_id = 'runwayml/
                        stable-diffusion-inpainting'
 8 pipe = AutoPipelineForInpainting.from_pretrained(
 9   model_id,
10   torch_dtype=torch.float16,
                                  # float16半精度の重みを使用
11   variant='fp16')              # fp16重みを使用
12 pipe.to('cuda')
13
14 # パイプラインを実行
15 init_image = load_image('init_image_01.JPG')
                                                  # 初期画像
16 mask_image = load_image('mask_01.jpg') # マスク画像
17 generator = torch.Generator(
       'cuda').manual_seed(138629319) # 乱数のシード値
18 image = pipe(prompt='Deep cracks on the road,
                        highly detailed', # プロンプト
19   image=init_image, mask_image=mask_image,
                     generator=generator).images[0]
20
21 # 画像を保存
22 image.save('sample_01.png')
```

● Pythonライブラリを読み込む（1行目～）

今回は，Diffusersライブラリから，パイプラインをAutoPipelineForInpaintingで作ります．インペインティングを行う入力画像は，load_imageで読み込みます．

● 画像生成パイプラインを設定する（6行目～）

このプロセスでは画像生成パイプラインを設定しています．モデルにstable-diffusion-inpaintingを利用し，重みとバリアントにfloat16の半精度を使用します．インスタンスを作り初期化します．また，このパイプラインは，CUDAコアで処理するように，pipe.to('cuda')を設定しています．

● パイプラインを実行する（14行目～）

このステップでは，画像生成パイプラインを実行します．15行目で画像修復に入力する初期画像，16行目に修復する部分を指定するのに使用するマスク画像のファイル名を設定しています．マスク画像の白い部分に画像がプロンプトに従った修復が行われます．

17行目に乱数シード値を使って，ジェネレータ・オブジェクトをCUDAコアにより生成して，変数generatorに格納します．18行目でパイプラインを実行します．実行する際にパラメータを設定します．最初にプロンプトを設定しており，日本語訳すると「道路の深い亀裂，非常に詳細」という内容にしています．その後，初期画像，マスク画像，ジェネレータを設定しています．生成された画像は，変数imageに出力しています．

● 生成された画像を保存する（21行目～）

22行目で生成された画像をpngファイルとして保存します．