

CycleGAN の学習に関する考察および画像生成モデル Stable Diffusion による画像編集の試み

板垣正敏¹

要旨

秋田県にある大湯環状列石の日時計状組石を題材に、石の風化や汚れを除去するための CycleGAN を用いた画像復元を試みてきたが、これまでのところ期待する成果は得られていない。

河原石での学習結果を見ると、背景を河原石の背景に近づける学習はできているように見えるが、石そのものは河原石に近づいているようには見えない。レプリカでの学習は、損失曲線を見る限りうまくいっているように見えるが、石の表面の違いが少ないため、その効果は確認できない。期待通りの学習ができていない理由として、次のようなことが考えられる。

①変化してほしい対象物である石の形状やテクスチャが一定でない、②河原石の場合、ドメイン A とドメイン B に石以外の共通項がない

CycleGAN の性質から、ドメイン B の画像が石の部分を除いてドメイン A に近いことが理想的である。つまり、レプリカの石の部分が風化や汚れのない新しい状態であるような画像が望ましい。しかしながら、そのような画像を用意することは困難である。

今回変化させたいのは画像の石の部分だけであることから、Mask あるいは Attention を使ったモデルでの学習を試すのが妥当であると考えます。

また、CycleGAN では変化させる対象を制御することは困難であるため、CycleGAN 以外の生成モデルとして、テキストでの指示により画像の変換を行うことができる Stable Diffusion での image2image 変換を試みた。

テキストによる指示の強さ $\text{strength} \cong 0.20$ では、石の形状が変化する、元の画像にないものが現れるなど「創作性」が強くなる。一方、 $\text{strength} \leq 0.15$ では、石の表面に大きな変化は見られない。

Stable Diffusion の image2image は、あくまでも入力画像を下敷きにした画像生成であり、どのように画像を変化させるかは、モデルがどのような画像・テキストを学習しているかに依存している。このため、「復元」と呼べるものではないと考える。

1 板垣中小企業診断士事務所