

# “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks”の学習報告

杉山 拓海<sup>1</sup>

**概要：** これまで深層学習を用いた画像から画像への変換学習では、ペアとなった訓練データセットが大量に用意できないような変換は学習が難しいという問題があった。この論文ではペアとなった訓練データセットを必要としない学習モデル Cycle-Consistent Generative Adversarial Networks(CycleGAN)を提案し、従来の学習手法との比較と、変換学習における汎用性の検証を行っている。比較では、ペア学習には劣るものの実用的なレベルの結果を示し、また汎用性検証では画像スタイル、模様変化、風景写真の季節など様々な変換実験を用いて汎用性を示している。

## 1. 学習構造

CycleGAN ではペア学習[1]で使用されていた図1左のような1対1に対応付けされたデータセットではなく、図1右のような画像群 X と画像群 Y という構造のデータセットを使用する。

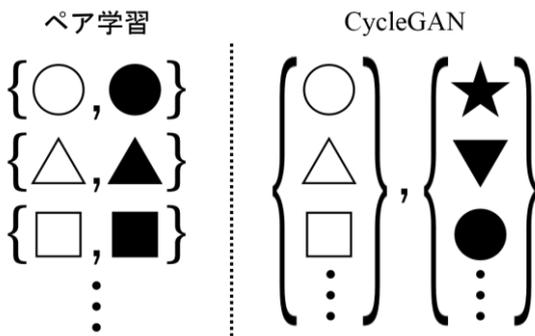


図1 訓練データセット

CycleGAN における学習の特徴は元画像から変換、復元を行った復元画像と元画像を比較するという部分にあり、図2は生成器 G で写真から地図を生成、生成器 F で地図から写真に戻したあと復元画像と元画像を比較という CycleGAN での学習の流れを示している。

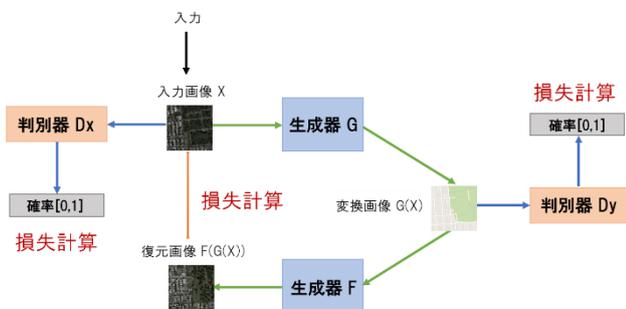


図2 CycleGAN の概略図

学習には損失 (loss) と呼ばれる値を使用する。CycleGAN で使用する全体の損失は以下の式で示される。

$$\begin{aligned} \mathcal{L}(G, F, D_X, D_Y) = & \mathcal{L}_{LSGAN}(G, D_Y, X, Y) \\ & + \mathcal{L}_{LSGAN}(F, D_X, Y, X) \\ & + \lambda \mathcal{L}_{cyc}(G, F) \end{aligned} \quad (1)$$

## 2. 学習結果

CycleGAN を用いた学習結果を図3に示す。高速道路など細かい部分まで正確に変換することは難しいが、緑地帯や道路の構造など全体的に特徴を捉えた変換は行えている。



図3 CycleGAN の学習結果 (写真⇄地図)

また、この論文ではペア学習で訓練データセットを用意することが困難だった変換についていくつか実験を行い、モデルの汎用性検証も行っている。変換の内容によってはうまく学習できないものもあるが、それでも幅広い種類の変換が期待できる。

## 参考文献

- [1] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint arXiv:1611.07004*, 2016

<sup>1</sup> 静岡大学 工学部 数理システム工学科  
〒432-8561 静岡県浜松市中区城北 3-5-1