

深層敵対的生成ネットワークを用いた マンガイラスト自動生成の試み

Trials of Generating Comic Illustrations Using Deep Generative Adversarial Networks

保住 純^{*1} 松尾 豊^{*1}
Jun Hozumi Yutaka Matsuo

^{*1}東京大学
The University of Tokyo

There have been lots of research of automatic generation of images using deep generative models, like Generative Adversarial Networks (GANs.) However most of these attempts are using colored photographs for training models and few researches on generating illustration are using deep generative models. In this paper, we show our results of several trials of generating comic illustration using recent GANs, trained by black and white illustrations in frames of "Tetsuwan Atom(Astro boy)" series written by Osamu Tezuka. We confirmed GANs can learn characteristics of comic illustrations, like lines and balloons, but we also found serious defects, one of which is indistinguishability of characters and backgrounds.

1. はじめに

これまでに人間の手によって数多くのマンガが制作されてきた。マンガを制作することは、キャラクターやストーリーを創造し、それを適切なコマ割りや様々な技法を用いることで画像として表現する技術であり、それは人工知能にとって難易度が高いタスクであると考えられる。そこで、本研究ではマンガにおける1コマ分の画像^{*1}をマンガイラストと定義し、それを自動生成することに焦点を置く。

近年、深層学習を用いて自動的に画像を生成する研究が数多く行われてきた。その際に用いられるモデルの一つに敵対的生成ネットワーク(Generative Adversarial Network, GAN)があり、高品質な画像を生成するための様々な手法が提案されてきた([Alec 15][Ishaan 17][Naveen 17])。ただし、これら研究の多くが検証に用いるデータセットは、MNISTのような非常に単純なものか、CIFAR-10やCelebAなどイラストに比べて情報量が多い写真画像を対象としたものとなっている。そのため、主にモノクロの線画によって構成され、かつ様々なシーンが描画されるマンガイラストには、これらの手法が適切とは言えない可能性がある。

一方で、イラストをGANによって自動生成する研究[Yanghua 17]や、マンガを対象として人工知能を活用する研究も従来より数多く行われている[松下 17]。しかし、マンガイラストを深層学習によって直接生成しようとする試みは、あまり行われていない。その原因の一つに、それをを行う上で学習に必要なデータセットが入手しづらいことが考えられる。もし仮にマンガイラストが収集できたとしても、学習に用いるマンガイラストの作画者が異なる場合、その表現方法の差異によって学習が困難になる可能性もある。

以上を踏まえ、本研究では深層敵対的生成ネットワークを用いたマンガイラストの自動生成を試みる。学習に用いるデータセットに、手塚治虫のマンガ「鉄腕アトム」(本編)全18巻に含まれる各コマを切り出した画像を用いる。使用する層の数や損失関数を変更することで複数回実験を行い、その結果を考察する。本研究にて得られる示唆は、人工知能によるマンガの自動生成を研究していく上で有益なものになると考えられる。

連絡先: 保住 純 (hozumi@weblab.t.u-tokyo.ac.jp)

^{*1} ここではフキダシ内の文章等を考慮しない。

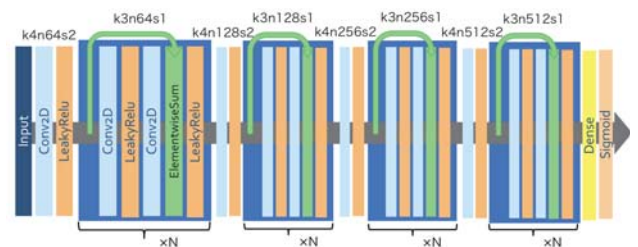


図 1: 識別ネットワークの構造

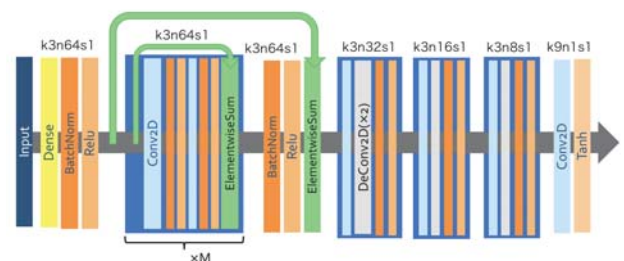


図 2: 生成ネットワークの構造

2. 手法

2.1 ネットワーク構造

本研究で用いるネットワークの構造を、図1と図2に記す。この構造を考える上で[Yanghua 17]を参考にした。図中の k はカーネル数、 n は特徴マップ数、 s はストライドを表す。また、図1での $\times N$ 、図2での $\times M$ の表記はその部分の繰り返し回数を示し、その値を変更することでネットワークの深さが変化する。

2.2 損失関数

本研究では比較のために、複数種類の損失関数(DCGAN, WGAN-GP, DRAGAN)を採用する。どちらも生成ネットワーク側の損失関数は共通で $L_G = \mathbb{E}[\log G(z)]$ であるが、識別ネットワーク側の損失関数 L_D がそれぞれ異なる。

- DCGAN [Alec 15]

$$L_D = \mathbb{E}[\log G(z)] + \mathbb{E}[\log(1 - G(z))]$$

- WGAN-GP [Ishaan 17]

$$L_D = \mathbb{E}[\log G(z)] + \lambda \mathbb{E}[|\nabla(D(\alpha x - (1 - \alpha G(z))))| - 1]^2]$$

- DRAGAN [Naveen 17]

$$L_D = \mathbb{E}[\log G(z)] + \lambda \mathbb{E}[|\nabla(D(\alpha x - (1 - \alpha x_p)))| - 1]^2]$$

3. 実験

3.1 設定

損失関数のハイパーパラメータとして $\lambda = 0.25$, 損失関数の最適化手法には Adam を採用し, $\alpha = 0.0002, \beta_1 = 0.5$ とした. ノイズ z の次元数は 100 とし, すべての重みは平均 0, 標準偏差 0.02 のガウス分布によって初期化する. 訓練時のバッチサイズは 64, イテレーション数は 40000 に設定した.

学習に用いるデータセットには, マンガ「鉄腕アトム」(本編) 全 18 巻に含まれているコマを手動で長方形に切り出した画像, 計 26428 枚を用いる. また, 各コマはサイズが 128×128 に統一されるよう拡大縮小を行った.

3.2 実験結果

実験結果を表 1 に記す. 結果欄における「○」は異なるノイズで異なる画像が出力された, 「×」はモード崩壊が発生しすべて同様の模様状の画像が生成されたことを示す. また, 各設定の中から筆者の目視判断によって質が最も良い画像を出力できた例を図 3 に記す.

N	M	損失関数	結果
1	4	DCGAN	○ (図 3)
1	4	WGAN-GP	○
1	4	DRAGAN	×
2	8	DCGAN	○
2	8	WGAN-GP	×
2	8	DRAGAN	×

4. 考察

今回の実験では, 層の数を少なくし, 損失関数に [Alec 15] のものを用いるという, よりシンプルな構造の GAN を採用するほうが良い出力結果を得られた. また, WGAN-GP や DRAGAN の損失関数の状況を見ると識別ネットワークの損失関数の最適化が生成ネットワークに比べて急激に進んでいたことが分かった. これはハイパーパラメータや各ネットワークの更新頻度を調整することで改善できる余地があると考えられる.

また, 生成された画像を調べると, コマの外枠を出力できているとともに, 細線で区切られたフキダシのような形状と, その中の文字に相当するであろう模様が生成できていた. 今後も生成ネットワークのアップスケーリングの工夫や画像出力後のノイズ除去処理を行うことによって, より高品質なイラストが出力できる可能性があると考えられる.

しかし, 生成画像を見ると背景とキャラの区別がつけられていない. そのため, もし次のステップとして「アトムが宇宙空間を飛行する」といった条件付き画像生成を考えるならば, ネットワーク構造の変更も含めた多方面での検討を行う必要があると考えられる.

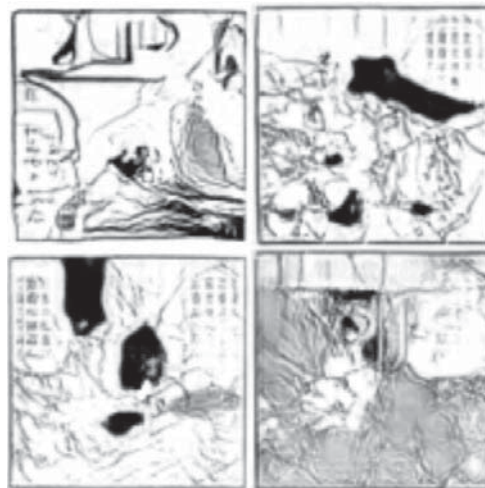


図 3: 本手法によって生成されるマンガイラストの例 (4 枚分)

5. 結論

本研究では, 深層敵対的生成ネットワークを用いてマンガイラストの自動生成を試みた. 使用する層の数や損失関数を変更することで複数回実験を行い, その結果を考察した. 今回はシンプルな構造の GAN ほど良い出力結果を得たが, それぞれの手法には様々な改良の余地があると考えられる.

しかし, マンガ自体を自動生成するためには, 他にも数多くの課題が残されている. その中にはストーリー生成やキャラクターと設定の創造といった, 画像処理でないタスクも多く存在する. それでも, 本研究が人工知能によるマンガの自動生成に向けた研究において少しでも参考になれば幸いである.

謝辞

学習用データの準備に際し, フィールド株式会社にご協力いただきました. この場を借りてお礼申し上げます.

参考文献

- [松下 17] 松下 光範, 山西 良典, 松井 勇佑, 岩田 基, 上野 未貴, 西原 陽子, 中村 聡史. 私のブックマーク「コミック工学 (Comic Computing)」, 人工知能 32 巻 6 号 (2017).
- [Yanghua 17] Jin, Yanghua, et al. "Towards the Automatic Anime Characters Creation with Generative Adversarial Networks." arXiv preprint arXiv:1708.05509 (2017).
- [Alec 15] Radford, Alec, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks." arXiv preprint arXiv:1511.06434 (2015).
- [Ishaan 17] Gulrajani, Ishaan, et al. "Improved training of wasserstein gans." Advances in Neural Information Processing Systems (2017).
- [Naveen 17] Kodali, Naveen, et al. "On Convergence and Stability of GANs" arXiv preprint arXiv:1705.07215 (2017).