



## 感性を考慮した画像生成

荒川 薫 (明治大学)

kara@meiji.ac.jp

### 1. 感性を考慮した画像生成とは

人工知能 (AI) を使って画像を生成する際、画像に対する人の感性を考慮する仕組みを取り入れ、人の好みや理想を反映した画像を自動的に生成することができる。これが感性を考慮した画像生成である。画像は数理的にモデル化することが難しく、任意の画像を効果的に生成することは長い間の課題であったが、21 世紀に入ってからの AI 技術により、それが可能となった。特に、2010 年代、Hinton らが階層形ニューラルネットワークの多層化、すなわち、深層学習を発表し、それまで実現が困難だった高度な画像処理や画像認識を実現できるようになった<sup>(1)</sup>。更に、2014 年、Goodfellow らは、新しい深層学習の形態である Generative Adversarial Networks (GAN: 敵対的生成ネットワーク) を提案し、これにより自然な画像を人工的に生成できるようになった<sup>(2)</sup>。GAN は二つの深層ニューラルネットワークを互いに競わせて、自然な画像を生成する。この基本構成を図 1 に示す。GAN は、基本的には、生成器 (Generator) と識別器 (Discriminator) からなり、これらは共に深層ニューラルネットワークである。GAN は、雑音などの入力から、本物らしいデータを生成しようと学習し、識別器は、生成器が生み出した偽物データを本物データから区別しようと学習する。GAN には多くの派生形があり、画像を入力して、それを別の見え方に変換する Cycle GAN<sup>(3)</sup>、pix2pix<sup>(4)</sup> などの GAN も提案されている。

GAN は、本物データとして集められた画像と類似する画像を生成することができるので、どのような画像を集めるかにより特性を設計できる。画像は、表示物体に関する客観的な情報を提供する一方で、「明るい感じ」「寂しい感じ」などの感性的な情報の提示も行う。このような感性的な情報には言葉でうまく表現できないようなものもある。そこで、人

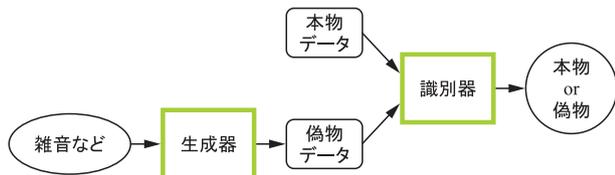


図 1 GAN の基本的な構成 雑音などを生成器に入力し、生成器は本物データとそっくりの偽物データを作ろうと学習する。識別器は本物データと偽物データを区別しようと学習する。両者が競い合うことで、生成器は本物と区別がつかないような偽物データを生成するようになる。

の感性を反映した画像を集めて本物データとすることで、その感性を反映した画像を生成することができる。更に、GAN に入力される雑音データを遺伝的アルゴリズムの染色体とみなし、対話形進化計算によって値を制御することにより、人の好みを反映させた画像を生成することができる。

### 2. 感性語の入力による人の感性の考慮

人の感性は言葉で表されることが多い。そこで、本物データとして、この感性語を反映した画像を集め GAN を学習させることで、その感性語を反映した画像を生成することができる。佐川らは、家具画像・感性語のデータセットから転移学習を行い、「シンプル」「美しい」などの感性語を ACGAN に入力することにより、感性語を反映したロゴマーク画像の生成を行う方法を提案した<sup>(5)</sup>。また、Takegishi らは、「柔らかな」「刺激的な」などの感性語を反映した形の生成を DCGAN により行い、それを基に、タイル形状をデザインする方を提案した<sup>(6)</sup>。図 2 に実際に生成されたタイル平面の例を示す。

### 3. スタイル変換

深層学習により画像の画風を他の画風に変換する方法が、2015 年に Gatys らにより提案された<sup>(7)</sup>。このように、画像の画風 (スタイル) を他のものに変換することをスタイル変換と呼ぶ。スタイル変換により、写真をそのコンテンツを変えずに自分の好みの画風、例えばゴッホの絵画風などに変換することが可能となる。Gatys らの方式は、畳込みニューラルネットワークを用いて実現され、出力画像が入力画像 (コンテンツ画像) のコンテンツを含み、画風の採用元の画像 (スタイル画像) と近い画風を持ち、更に滑らかな画像の体裁をなすように損失関数を設定して学習が行われる。

GAN を用いてスタイル変換を行うこともできる。Zhu らが 2017 年に発表した Cycle GAN は、図 3 のように二つのタイプの画像の集合 (ドメイン) 間で画像のスタイル変換を行う<sup>(8)</sup>。ここでは、ドメイン A の画像を生成器 G でドメイン B の画像になるように変換し、それがまた生成器 F で



(a) 落ち着いた (b) 柔らかな (c) 固い

図 2 感性語ごとに学習された DCGAN を用いて作られたタイル形状と、それを敷き詰めたタイル平面。ここでは DCGAN により生成された形状に、平面を敷き詰め可能となるための制約が加えられている。

本会ハンドブック「知識の森」  
[https://www.ieice-hbkb.org/portal/doc\\_index.html](https://www.ieice-hbkb.org/portal/doc_index.html)

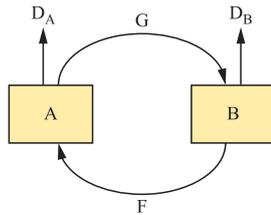


図3 CycleGANの構成 ドメインAの画像を生成器GによりドメインBの画像に変換し、それを生成器FによりドメインAの元の画像に近くなるように学習する。D<sub>A</sub>はAの本物と偽物、D<sub>B</sub>はBの本物と偽物を区別するよう学習される識別器。

メインAの元の画像に近付くように学習が行われる。この方式では、一方のドメインを写真、他方を好みの絵画とすると、写真をそのコンテンツを変えないで、好みの絵画風に変換することができる。

#### 4. 対話形進化計算によるGANの潜在ベクトルの制御

対話形進化計算は、システムを設計する際、このシステムパラメータを生物の染色体に見立て、遺伝的アルゴリズムなどの進化計算により最適化を行う<sup>(9)</sup>。ここで、人の主観評価を用いてシステム出力の評価を行う。この構成図を図4に示す。すなわち、対話形進化計算は初期世代の個体(染色体)としてランダムな数値系列を複数個発生させ、これらの個体を使って画像などのデータを生成し、人がそれら进行评估する。これらの生成物の中に、人が十分満足できるものがあればそれを選択して終了するが、そうでなければ、比較的评价の高い個体を何個か選ぶ(選択)。次に、選択された個体と似ている個体と、似ていない個体を新たに生成し、これらと選択された個体を合わせて次世代の集団とする。ここで、選択された個体に似ている個体を作ることを交差、異なる個体を作ることを突然変異と呼ぶ。次世代においては、新しい個体を使ってデータを生成し、評価、選択、交差、突然変異を繰り返す、最終的に満足できるデータが得られれば終了する。

GANは、乱数列を潜在ベクトルとして用い、それに基づいて多様な画像を生成する。この潜在ベクトルは、生成される画像と関連付けられるが、特定の画像を生成するために乱数の値を制御することは行われない。そこで、この潜在ベクトルを対話形進化計算で扱われる染色体とみなすと、GANが生成するデータに対して対話形進化計算を適用することができる。これは、GANが生成する多数の画像等に対し、人が主観的に高く評価するものを探索していくことに相当する。

このように、対話形進化計算をGANに導入して潜在変数を制御する方法として、Bontragerは、DCGANに対話形進化計算を導入し、利用者の好みや主観評価を考慮した物体のデザインや顔画像の生成を実現できることを示した<sup>(10)</sup>。しかし、一般的な物体のデザインは自由度が高く、DCGANでは所望のデザインが出現しづらい傾向がある。そこで、Xinらは、所望のデザインの輪郭画像を入力し、この輪郭の中身をGANと対話形進化計算の融合により生成する方式を提案した<sup>(11)</sup>。ここでは、BicycleGAN<sup>(12)</sup>を用いる。BicycleGANはpix2pixのように、与えられた輪郭を有する画像を生成するGANであるが、乱数列を潜在変数として用いることで、画像に多様性を持たせることができる。図5はこの方式で実際に生成された靴のデザインの例を示したものである。

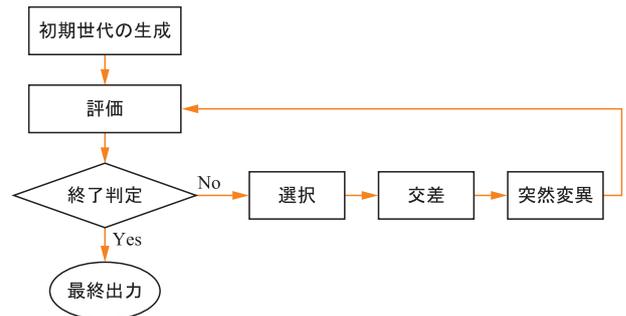


図4 対話形進化計算の流れ 評価は人が主観に基づいて行う。

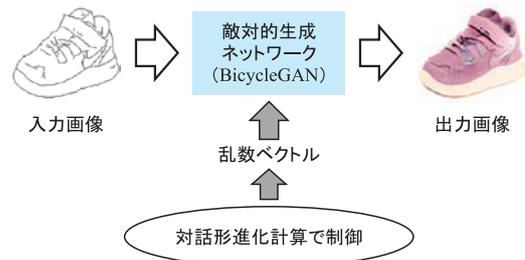


図5 対話形進化計算を導入したGANにより生成された靴のデザインの例 輪郭画像を入力し、色や素材感をデザイン。対話形進化計算によりBicycleGANへの入力雑音(乱数ベクトル)を制御。

#### 5. 解像度ごとのスタイル情報に対する対話形進化計算の導入

高精細な画像を生成するGANとして、低解像度画像から学習し、解像度を上げていくPGGANが2018年に発表されたが<sup>(13)</sup>、これにスタイル変換を取り入れたのがStyleGANである<sup>(14)</sup>。このStyleGANはランダムな潜在ベクトルを、スタイル情報を表す潜在空間に変換し、それを解像度ごとに深層ニューラルネットワークの各畳込み層に入れ込む。この潜在空間におけるベクトルもランダムな数値系列として表されるので、複数の生成画像の潜在空間を同じ解像度同士で混ぜたり入れ替えたりすることで、画像のスタイルを考慮しながら交差を実現することができる。Tejeda-Ocampoらは、このような解像度ごとの潜在空間の統合を行う対話形進化計算により、絵画を効果的に生成する方法を提案した<sup>(15)</sup>。また、Liaoらは、潜在空間の低・中・高の解像度領域が、映画ポスターのタイトル、人物の構成、全体の色味のスタイルに対応することを明らかにし、解像度ごとの潜在空間の入替えを行うことで利用者の好みの映画ポスターのレイアウトを生成する方式を提案した<sup>(16)</sup>。

#### 6. 人の感性に対するAIの貢献

AI、特に深層学習の躍進により、人が本物の写真や絵画とみなすような画像をコンピュータ上で生成できるようになった。また、人が、好ましいと思う画像そのものを入力したり、その画風を言葉で表して入力することで、AIは人の感性とそれを反映した画像の間をつなぐ数値モデルを実現し、人の感性を考慮した画像生成が可能となった。今後もAIが、人の精神的側面を更に充実させることが期待される。

## 文 献

- (1) A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems* 25, pp. 1097-1105, 2012.
- (2) I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems* 27, 2014.
- (3) J-Y. Zhu, T. Park, P. Isola, and A.A. Efros, "Unpaired Image-to-Image translation using cycle-consistent adversarial networks," arXiv:1703.10593 [cs. CV], 2017.
- (4) P. Isola, J-Y. Zhu, T. Zhou, and A.A. Efros, "Image-to-image translation with conditional adversarial networks," arXiv:1611.07004 [cs. CV], 2017.
- (5) 佐川友里香, 萩原将文, "感性語を考慮した敵対的生成ネットワークを用いたロゴマーク生成," *日本感性工学会論文誌*, vol. 18, no. 3, pp. 215-222, 2019.
- (6) M. Takegishi and K. Arakawa, "Interactive evolutionary design method of tile shape by GAN reflecting adjectives," 2021 ISPACS, Nov. 2021.  
DOI:10.1109/ISPACSS51563.2021.9651018
- (7) L.A. Gatys, A.S. Ecker, and M. Bethge, "A neural algorithm of artistic style," arXiv:1508.06576 [cs. CV], Aug. 2015.
- (8) J.Y. Zhu, T. Park, P. Isola, and A.A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," *Proc. IEEE Int. Conf. Computer Vision (ICCV)* 2017, pp. 2223-2232, 2017.
- (9) H. Takagi, "Interactive evolutionary computation : fusion of the capabilities of EC optimization and human evaluation," *Proc. IEEE*, vol. 89, pp. 1275-1296, 2001.
- (10) P. Bontrager, et al., *Deep interactive evolution Int. Conf. Computational Intelligence in Music, Sound, Art and Design*, Springer, Cham, 2018.
- (11) C. Xin and K. Arakawa, "Object design system by interactive evolutionary computation using GAN with contour images," *KES-HCIS 2021, Smart Innovation, Systems and Technologies*, vol 244. Springer, Singapore, 2021.  
[https://doi.org/10.1007/978-981-16-3264-8\\_7](https://doi.org/10.1007/978-981-16-3264-8_7)
- (12) J. Zhu, et al., "Toward multimodal image-to-image translation," *Advances in neural information processing systems*, 2017.
- (13) T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," arXiv:1710.10196 [cs. NE].
- (14) T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," arXiv:1812.04948 [cs. NE].
- (15) C.T. Ocampo, et al., "Improving deep interactive evolution with a stylebased generator for artistic expression and creative exploration," *Entropy*, 23, 11, 2020.  
<https://doi.org/10.3390/e23010011>
- (16) S. Liao and K. Arakawa, "Interactive poster design system for movies with StyleGAN," 2021 Int. Symp. Intelligent Signal Processing and Communication Systems (ISPACS), pp. 1-2, doi:10.1109/ISPACSS51563.2021.9651023, Hualien City, Taiwan, 2021.

(2023年10月19日受付)

