

1. はじめに

深層学習による生成モデルの1つである Generative Adversarial Networks (GAN) は、敵対的な学習により画像を生成することができる。Conditional GAN (CGAN) では、クラスラベル等の条件を入力することで条件を満たした画像生成が可能である。しかし、実画像と比較した際に生成画像が不自然であるという問題がある。そこで本研究では、重み付き条件を CGAN に導入した顔画像生成法を提案する。

2. Generative Adversarial Networks

GAN[1] は、ノイズベクトル $z \in [-1, 1]$ を入力して画像を生成する Generator と、実画像か生成画像かを判別する Discriminator で構成されている。Generator と Discriminator を敵対的に学習させることで、データセットに存在しない画像を生成できる。CGAN は、クラスラベル等の条件 $y \in \{0, 1\}$ をネットワークに入力することで、条件を考慮した画像の生成が可能である。さらに、畳み込み層を導入することで、より高品質な画像を生成できる Deep Convolutional GAN (DCGAN)[2] や Conditional DCGAN, 階層的な生成で高解像度な画像となる Progressive Growing GAN (PGGAN) がある。しかし、CGAN では条件を 0 と 1 の 2 値で表現して与えるため中間的な表現ができない。そのため、自然な顔画像の生成が難しいという問題がある。

3. 提案手法

本研究では、Generator に重み付きで条件を入力し、Discriminator で条件を認識する GAN を提案する。図1に提案手法のネットワーク構造を示す。ここで、図1の x は実画像、 $G(z|y)$ は Generator の出力、 $D(x|y)$ または $D(G(z|y)|y)$ は Discriminator の真偽に対する出力、 $C(G(z|y)|y)$ は Discriminator の認識に関する出力である。各ネットワークの学習について以下に述べる。

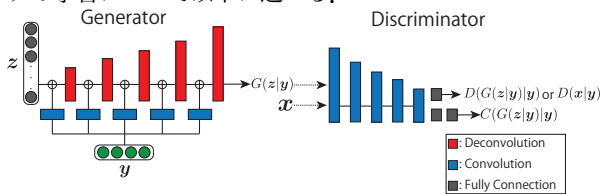


図1：提案手法のネットワーク構造

3.1. Generator の学習

Generator では、各 Deconvolution 層にて入力する条件ベクトル y の畳み込み処理を行う。これにより、2 値の条件ベクトル y を重み付きの連続値として表現できる。また、重みを学習により各層毎に決定することで、段階的に条件を反映した画像生成が期待できる。DCGAN に提案手法を導入する場合は、図1となる。また、PGGAN に提案手法を導入する場合は、階層的に条件ベクトル y の畳み込み層を追加する。

3.2. Discriminator の学習

Discriminator に実画像 x もしくは生成画像 $G(z|y)$ が入力された際に、同時に条件ベクトル y を入力する。さらに提案手法の Discriminator は、与えられた条件を認識するようにマルチタスク化する。これにより、条件をより反映した Generator の敵対的学習が期待できる。提案手法の目的関数は式 (1) となる。

$$L = \log D(x|y) + \log[1 - D(G(z|y)|y)] + C(G(z|y)|y) \quad (1)$$

4. 評価実験

本実験では、CelebA Dataset を用いて顔画像の生成をする。CelebA Dataset は、202,599 枚の顔画像から構成されており、40 種類の顔属性が付与されている。本実験で用いる学習サンプル数は 100,000 枚、Male (性別)、Eyeglasses (メガネ)、Smiling (表情)、Goatee (ヒゲ)、Bangs (前髪) の 5 つの顔属性を 5 次元の条件ベクトル y として入力す

る。 z は、100 次元のノイズベクトルとする。生成画像の客観評価には Inception Score を用いる。また、主観評価も同様に行う。

4.1. 実験結果

提案手法で顔画像を生成した結果を図2に示す。異なる条件を指定した場合でも、高品質な顔画像を生成できていることがわかる。

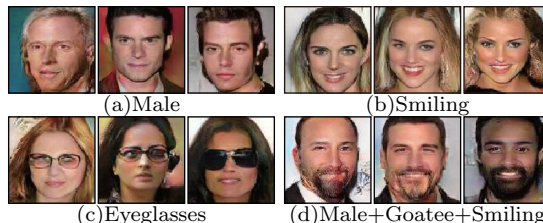


図2：提案手法による顔画像生成例

従来手法 (Conditional DCGAN), 提案手法 (DCGAN), 提案手法 (PGGAN) の評価結果を表1に示す。表1より、Inception Score, 主観評価ともに提案手法 (PGGAN) が最も高い。

表1：評価結果

手法	Inception Score	主観評価 (12人)
実画像	4.15	-
従来手法	2.05	39.3
重み付き DCGAN	2.16	35.7
重み付き PGGAN	3.00	75

4.2. 考察

Generator の各層での顔属性の寄与率 C_t を式 (2) により求める。ここで、 M は属性数、 N は重みフィルタ数、 W は重みフィルタ、 t は寄与率を算出する属性である。

$$C_t = \frac{1}{N} \sum_{n=1}^N \frac{\sqrt{(W_{t,n})^2}}{\sum_{m=1}^M \sqrt{(W_{m,n})^2}} \quad (2)$$

各層の顔属性の寄与率を図3に示す。1層目では Male が他の属性と比較すると最も高い寄与率であり、性別に着目して生成している。2層目では、Bangs (前髪) の寄与率が高い。3層目以降では各属性の寄与率に大きな差はない。以上より、早い段階では性別と髪の色に注目し、後段では全ての条件を満たすように顔画像を生成しているといえる。

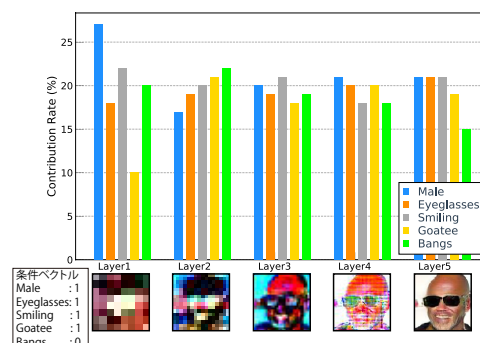


図3：各層の顔属性の寄与率と中間生成画像

5. おわりに

本研究では、Generator の各層に重みを導入した条件付き GAN を提案し、Inception Score と主観評価により、提案手法の有効性を確認した。今後は、さらなる生成画像の高解像度化に取り組む。

参考文献

[1] I. Goodfellow, et al., "Generative Adversarial Nets", NIPS, pp. 2672-2680, 2014.
 [2] A. Radford, et al., "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", ICLR, 2016