

1. はじめに

Generative Adversarial Networks (GANs) は Generator および Discriminator を敵対的に学習することで、実在しない画像を生成する手法である。Generator は、潜在変数を用いて画像を生成し、生成画像が Discriminator により実画像と識別されるように学習する。Discriminator は、実画像と生成画像を識別できるように学習する。GANs の学習時、Generator の生成画像に偏りが生じると、Discriminator は実画像と生成画像を簡単に識別でき、モード崩壊を起こすことがある。この問題を解決するために、複数の Generator を用いることで、モード崩壊の発生を抑制する手法が提案されている [2]。しかし、この手法は Generator が独立して画像を生成するため、生成画像に偏りが生じる問題は解決していない。そこで、本研究では生成画像の偏りを抑制するために、複数の Generator が知識を転移しながら共同学習して画像を生成する手法を提案する。

2. 関連研究

従来手法である Deep Convolutional GANs および Multi-Agent Diverse GANs について述べる。

Deep Convolutional GANs (DCGANs) [1] は、畳み込み層を用いることで生成画像の質を向上させた GANs の手法である。実画像として正面の顔画像のみのデータ等の類似したデータのみ扱う場合においては、実画像と区別ができない鮮明な画像を生成できる。学習する際にモード崩壊が発生する場合が多く、安定した学習が困難である。

Multi-Agent Diverse GANs (MAD-GANs) [2] は、複数の Generator および 1 つの Discriminator で構成された GANs の手法である。Discriminator は実画像と生成画像を識別するのみならず、どの Generator から生成された生成画像であるかを識別するように学習する。MAD-GANs は、1 つ Generator が生成する画像に偏りが生じた場合、他の Generator は偏った画像とは異なる画像を生成しており、Discriminator は実画像と生成画像を簡単には識別できない。これにより、MAD-GANs を用いることでモード崩壊の発生を抑制することが可能となる。しかし、生成画像の偏りが生じること自体を防ぐことはできない。

3. 提案手法

本研究では複数の Generator を用いて学習を行い、Generator が互いの知識を転移しながら共同学習して画像生成を行う GANs の手法を提案する。

3.1 GANs の共同学習

提案手法のネットワーク構造を図 1 に示す。学習に使用する Generator の数を k とすると、各 Generator は G_1, \dots, G_k で表される。各 Generator は、潜在変数 $\mathbf{z} \sim P_z$ を用いて生成画像を生成する。そのため、 k 枚の生成画像が得られる。また、各 Generator の特徴マップまたは生成画像を知識転移グラフを用いて転移させることで、Generator 同士での知識の伝達を可能にしている。Discriminator は D で表され、実画像と生成画像を識別するのみならず、どの Generator から生成された生成画像であるかを識別する。そのため、最終層の出力値のユニットの数は $k+1$ 個となっており、ソフトマックス関数を用いてクラススコア d_1, \dots, d_{k+1} を出力する。スコア d_1, \dots, d_k は各 Generator から生成された生成画像である確率を表し、スコア d_{k+1} は実画像である確率を表す。ネットワーク内部には、学習の安定化を図るため、Batch Normalization が用いられる場合が多い。しかし、GANs における学習の安定化には Batch Normalization は不十分である。そこで、本手法では Discriminator に Spectral Normalization (SN) を導入する。これにより、損失関数に変更を加えることなくモード崩壊を抑制し、GANs における学習の安定化を図ることが可能となる。

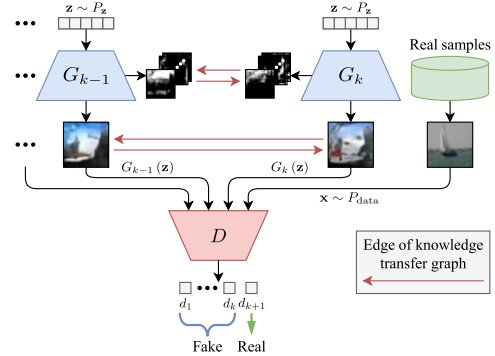


図 1: 提案手法のネットワーク構造

3.2 知識転移グラフによる最適化

提案手法は、知識転移グラフを導入することで、各 Generator が互いに知識転移しながら学習する。3 つの Generator を用いた場合における知識転移グラフを導入した提案手法を図 2 に示す。学習に使用する Generator の数を k とすると、各 Generator は G_1, \dots, G_k 、Discriminator は D となり、ノードとして表す。各 Generator のノード間には 2 つのエッジを定義し、それぞれ異なる向きの有向グラフで表す。各エッジの向きは、学習時に勾配の情報が伝わる方向を表している。各 Generator のノード間のエッジは Generator 同士の知識の伝達を表し、 i 番目の Generator のエッジ $L_{D,i}$ は i 番目の Generator が生成した生成画像が Discriminator に実画像と判断されたかどうかの情報の伝達を表す。また、エッジ $L_{adversarial}$ は、Discriminator の識別結果が正しいかどうかの情報の伝達を表す。各エッジには個別の損失関数を定義し、各 Generator の方向を向いたエッジの損失関数にはゲート関数を導入することで、伝播される損失を制御する。本手法では、ゲート関数として Through Gate, Cutoff Gate, Negative linear Gate および Positive linear Gate を用いる。Through Gate は、入力されたサンプルごとの損失をそのまま通す関数である。Cutoff Gate は損失を 0 とし、損失の計算を行わない関数であり、任意のエッジを切断することができる。Negative linear Gate は学習が進むにつれて損失を減少させる関数であり、学習の序盤は Through Gate、終盤は Cutoff Gate と同じ処理となる。Positive linear Gate は学習が進むにつれて損失を増加させる関数であり、学習の序盤は Cutoff Gate、終盤は Through Gate と同じ処理となる。

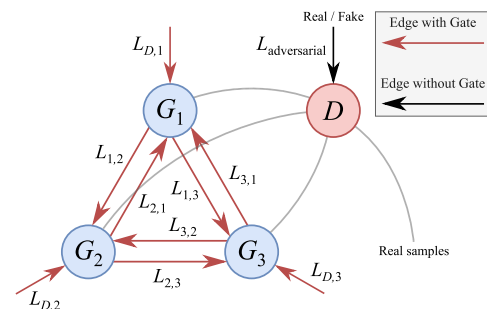


図 2: 複数の Generator における知識転移グラフ ($k=3$)

3.3 学習方法

提案手法では、Discriminator の学習および Generator の学習を交互に繰り返すことで GANs の学習を行う。また、Discriminator の学習は MAD-GANs と同様である。

Generator の学習は、各 Generator の生成画像が Discriminator により実画像と識別されるように行う。また、提案手法では知識転移グラフに応じて、各 Generator が互いの知識を転移しながら学習する。各 Generator のノード間のエッジは Generator 同士の知識の伝達を表し、 i 番目の Generator のエッジ $L_{D,i}$ は i 番目の Generator が生成

した生成画像が Discriminator に実画像と判断されたかどうかの情報の伝達を表す。Generator のノード間のエッジで、 t 番目の Generator の知識を s 番目の Generator に転移する際の損失関数を $L_{t,s}$ とし、 i 番目の Generator が生成した生成画像が Discriminator に実画像と判断されたかどうかを表す損失関数を $L_{D,i}$ とする。また、各エッジにおけるゲート関数は $Gate(\cdot)$ とする。Generator の学習では、各 Generator に入力される潜在変数 $\mathbf{z} \sim P_z$ は同一とする。最終的な各 Generator の損失関数はエッジから得られた $L_{t,s}$ および $L_{D,i}$ を総和して算出する。使用する Generator の数を k とすると、Discriminator に対して i 番目の Generator が生成した生成画像を入力した際の識別結果は $\mathbf{d}^{(i)} = d_1^{(i)}, \dots, d_{k+1}^{(i)}$ で表される。また、正解ラベルは $\mathbf{t}^s = t_1^s, \dots, t_{k+1}^s$ で表され、 t_{k+1}^s が 1 でその他が 0 のベクトルとなる。このとき、損失関数 $L_{D,i}$ は式 (1) となる。

$$L_{D,i} = Gate \left(- \sum_{j=1}^{k+1} t_j^s \log_e d_j^{(i)} \right) \quad (1)$$

提案手法では、特徴マップを転移することにより、知識転移を行う。特徴マップのチャンネル番号を $c = 1, \dots, C$ 、特徴マップの層の位置を $m = 1, \dots, M$ 、特徴マップのベクトル番号を $n = 1, \dots, N$ とすると、 i 番目の Generator の特徴マップは $F_{i,m,c,n}$ で表される。このとき、ハイパーパラメータを α とすると、損失関数 $L_{t,s}$ は式 (2) および式 (3) となる。

$$Q_i^{(m,n)} = \frac{1}{C} \sum_{c=1}^C F_{i,m,c,n}^2 \quad (2)$$

$$L_{t,s} = Gate \left(\frac{\alpha}{MN} \sum_{m=1}^M \sum_{n=1}^N \left(\frac{Q_t^{(m,n)}}{\|Q_t^{(m,n)}\|_2} - \frac{Q_s^{(m,n)}}{\|Q_s^{(m,n)}\|_2} \right)^2 \right) \quad (3)$$

4. 評価実験

従来手法との比較により、提案手法の有効性を調査し、最適な知識の転移方法を探索する。実験には、10 クラスの一般物体のカラー画像で構成されている CIFAR-10 データセットを用い、学習用のサンプルである 50,000 枚を 64×64 にリサイズして使用する。本実験では、ベースの GANs の手法として DCGANs を用いる。また、知識転移グラフにおけるゲート関数の探索回数を 1,500 とし、各 Generator における生成画像の Inception Score (IS) を最大化するゲート関数の組み合わせを探索する。学習時のバッチサイズを 256 とし、学習回数を 50、各 Generator に入力する潜在変数の次元数を 100、ハイパーパラメータ α を 1,000 とする。生成画像の質を評価する指標として IS および Fréchet Inception Distance (FID) を用いる。また、IS は高い値、FID は低い値になるほど生成画像の質が高いことを表している。

4.1 提案手法の有効性の調査

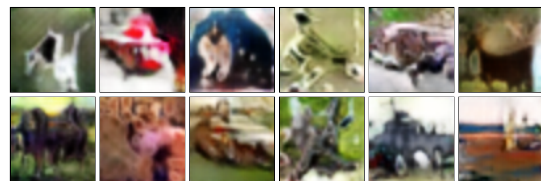
本実験では、DCGANs および MAD-GANs、提案手法における生成画像の質を比較する。DCGANs は SN の導入の有無による 2 パターンを比較対象とする。MAD-GANs および提案手法は SN を導入し、Generator の数を 3 とする。学習回数を 200 Epoch とし、学習終了時のモデルパラメータを用いて評価を行う。各手法の IS および FID を表 1 に示す。表 1 より、提案手法は IS が最も高い値となり、最も生成画像の質が高いことが確認できる。FID は SN を導入した DCGANs が最も低く、提案手法も同様に低い値である。GAN は学習が進むと、IS と FID がトレードオフの関係にあることが知られている。本実験における提案手法は、各 Generator における生成画像の IS を最適化対象とし、最大化するゲート関数の組み合わせを探索した。提案手法は最適化対象を FID とし、最小化するゲート関数の組み合わせの探索により、FID に特化した手法も実現できる。

4.2 生成画像の可視化

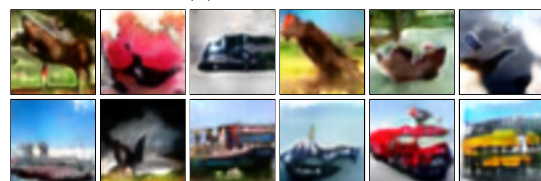
MAD-GANs および提案手法に同一の潜在変数を入力し、生成した画像を比較する。生成画像の可視化結果を図 3 に示す。図 3 の結果から、提案手法は実画像に近い画像が生成できていることが確認できる。

表 1: IS および FID の比較結果

	SN	IS \uparrow	FID \downarrow
DCGANs		4.57 ± 0.03	25.79
	✓	4.77 ± 0.05	21.17
MAD-GANs	✓	4.77 ± 0.03	28.76
提案手法	✓	5.37 ± 0.09	23.77
実画像	—	8.70 ± 0.14	—



(a) MAD-GANs



(b) 提案手法

図 3: 生成画像の可視化結果

4.3 知識転移グラフの最適化

提案手法において、知識転移グラフを用いたゲート関数の探索を 1,500 回行い、知識の伝達方法を最適化した。最適化対象である各 Generator における生成画像の IS が最大の値となったグラフを図 4 に示す。赤色のエッジが Positive linear Gate、灰色のエッジが Through Gate を表している。 G_1 は、他の Generator から知識を転移されない G_2 と比較して IS の値が高く、知識の転移は生成画像の質の向上に有効であるといえる。また、Generator 間の知識の転移は学習初期は不要であり、学習が進むにつれて転移すると良いことや、 G_3 は G_1 のバッファのような役割を果たすことで、 G_1 の学習の安定化に寄与していることが確認できる。

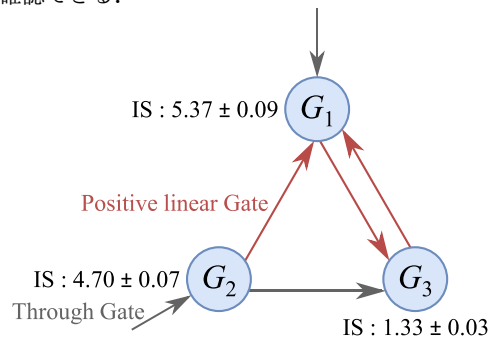


図 4: 知識転移グラフの最適化結果

5. おわりに

本研究では複数の Generator および 1 つの Discriminator を用いて、それぞれの Generator で知識を転移しながら共同学習を行う手法を提案した。評価実験により提案手法の有効性を確認した。今後は複数の Discriminator を用いることで、アンサンブル効果により生成画像のさらなる質の向上を目指す。

参考文献

- [1] A. Radford, *et al.*, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks”, ICLR, 2016.
- [2] A. Ghosh, *et al.*, “Multi-Agent Diverse Generative Adversarial Networks”, CVPR, 2018.

研究業績

- [1] T. Tsukahara, *et al.*, “Collaborative learning of generative adversarial networks”, VISAPP, 2021.

(他 3 件)