

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 2/6/2019

専攻名 Department	情報通信	氏名 Name	大山優香	指導 教員 Advisor	渡辺 裕 ⑨
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1w152063-9 ^{CD}		
研究題目 Title	敵対的生成ネットワークによる和菓子生成画像の評価 Evaluation of Generated Japanese Sweets Images by Generative Adversarial Networks				

1. まえがき

和菓子は五感の芸術であると言われている。また同じ植物、同じ風景を題材にした和菓子であっても、使用した材料や製作する職人によってさまざまな表現が存在する。しかし、和菓子において独創的かつ分かり易いデザインを新たに考案することは容易でないと考えられる。

Generative Adversarial Networks (GAN)は2014年に Ian J. Goodfellow によって考案された生成モデルを訓練させるためのモデリング手法である [1]。近年 GAN を用いた Deep Learning による画像生成は様々な分野において行われている。GAN を用いることで、和菓子のデザインを考案することができると考えられる。本研究では GAN の発展形である Creative Adversarial Network (CAN), Semantic Decomposed GAN (SD-GAN)および WGAN-GP を用いて和菓子の画像を生成し、Convolutional Neural Networks (CNN) を用いて生成画像の評価を行う。

2. Creative Adversarial Networks

Creative Adversarial Networks (CAN)は2017年にラトガース大学の Ahmed Elgammal らが発表した GAN を基本とした画像生成技術で、新しい芸術を生み出すことができる [2]。芸術らしさを持つ画像かつ学習したスタイルへの適応度が低い画像を正解とすることによって、既存のスタイルから逸脱した芸術作品の画像の生成を行う。

3. Semantically Decomposed GAN

Semantically Decomposed GAN (SD-GAN)は2018年にカリフォルニア大学の Chris Donahue らが発表した GAN を基本とした画像生成技術である [3]。Siamese Network を用いることで1つの被写体のライティング、ポーズなどの異なる複数の画像を生成することができ、新しい被写体の画像を生成することもできる。

4. WGAN-GP

2017年にクーラント数理科学研究所の Martin Arjovsky らによって Wasserstein GAN (WGAN)が提案された [4]。しかし WGAN に用いられる weight clipping により学習が依然として不安定であった。

そこで WGAN を改善した WGAN-GP が2017年にモントリオール学習アルゴリズム研究所の Ishaan Gulrajani らによって提案された [5]。WGAN-GP はリップシツツ制約をみたすために weight clipping の代わりに gradient penalty を使用する。これにより安定した学習を行うことができる。

5. 提案手法

GANで和菓子画像のデータセットを用いて生成した画像が和菓子とみなせるかを、CNNを用いて評価する。まず、CNNに和菓子の画像と和菓子でない物体の画像を学習させる。識別精度の高いCNNを構築するために、学習には VGG16 モデルを用いて Fine Tuning を行う。全結合層とその一つ前の畳み込み層のブロックの重みを再学習し、それ以前の層の重みは固定する。

学習終了後、GANによって生成された画像を学習した CNN に入力し和菓子画像であるかどうかを判別する。複数枚生成画像を入力して和菓子画像と認識された画像の割合を求めることで正解率を算出する。

6. 実験

和菓子画像 11051 枚のデータセットを用いて CAN, SD-GAN, WGAN-GP を学習した。CAN による生成画像例を図 1, SD-GAN による生成画像例を図 2, WGAN-GP による生成画像例を図 3 に示す。

また、CNN を用いて入力された画像が和菓子画像か否かを判定する。CNN の学習は VGG16 モデルを用いて Fine Tuning を行った。学習後、評価用画像に

よる正解率は 98.91[%]であった。

CAN, SD-GAN, WGAN-GP によって生成されたそれぞれ 1000 枚の画像を CNN によって評価し、正解率を算出した。正解率を表 1 に示す。

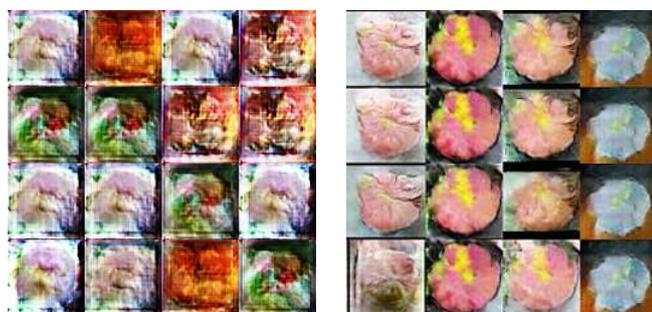


図 1 CAN の生成画像例 図 2 SD-GAN の生成画像例



図 3 WGAN-GP の生成画像例

表 1 生成画像の評価

	正解率[%]
CAN	7.9
SD-GAN	81.9
WGAN-GP	80.5

7. 考察

CNN の学習は、評価用画像による正解率が 98.91[%]であったことから成功したといえる。また、GAN による生成画像の評価結果で一番正解率が高いのは SD-GAN による生成画像であった。CNN による評価の結果からは、和菓子画像の生成に一番適しているのは SD-GAN であるといえる。

一方各 GAN の生成画像例を比較すると、CAN による生成画像と SD-GAN による生成画像は類似した画像が多く生成されている。WGAN-GP による生成画像は類似した画像が少なく、多様な画像が生成できた。目視による評価の結果からは、和菓子画像の生成に一番適しているのは WGAN-GP であるといえる。

CNN による評価は生成画像一枚ずつが和菓子で

あるかどうかを判別しているため、生成画像の類似性についての評価ができない。したがって目視による評価と異なる結果になった。CNN による評価において SD-GAN による生成画像の正解率は 81.9[%]、WGAN-GP による生成画像の正解率は 80.5[%]であり、WGAN-GP による生成画像も十分に高い正解率であるといえる。したがって和菓子画像の生成に一番適しているのは WGAN-GP であるといえる。

8. まとめ

本研究では3種類のGANによって和菓子画像を生成した。また Fine Tuning を用いて和菓子画像判別の CNN モデルを作成し、生成画像の正解率を評価した。

和菓子画像判別 CNN を用いて算出した GAN による生成画像の正解率は SD-GAN による生成画像が最も高かった。しかし、CAN および SD-GAN による生成画像は類似した画像が多かった。WGAN-GP による生成画像は多様性が高く、CNN による正解率も十分に高いため和菓子画像の生成に一番適しているのは WGAN-GP である。

本研究で作成した CNN モデルは入力画像 1 枚に対して和菓子であるか否かを判別するモデルであるため生成画像全体の多様性に対しての評価ができなかった。複数枚の画像の多様性に対して定量的な評価をする方法を検討すべきである。

参考文献

- [1] I. J. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville and Y. Bengio, "Generative Adversarial Nets," NIPS, pp.1-9, 2014.
- [2] A. Elgammal, B. Liu, M. Elhoseiny and M. Mazzone, "CAN: Creative Adversarial Networks Generating "Art" by Learning About Styles and Deviating from Style Norms," ICCV, pp.1-22, 2017.
- [3] C. Donahue, Z. C. Lipton, A. Balsubramani and J. McAuley, "Semantically Decomposing the Latent Spaces of Generative Adversarial Networks," arXiv preprint arXiv:1705.07904, 2017.
- [4] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein GAN," arXiv preprint arXiv:1701.07875, 2017.
- [5] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. Courville, "Improved Training of Wasserstein GANs," arXiv preprint arXiv: 1704.00028, 2017.

2018 年度 卒業論文

敵対的生成ネットワークによる和菓子生成画像の評価

Evaluation of Generated Japanese Sweets Images
by Generative Adversarial Networks

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報通信学科

1W152063-9

大山優香

目次

第1章	序論	1
1.1	研究の背景	1
1.2	本研究の目的	1
1.3	本論文の構成	1
第2章	CNN	3
2.1	まえがき	3
2.2	ニューラルネットワーク	3
2.3	Convolutional Neural Network	4
2.4	Fine Tuning	4
2.5	VGG16	5
2.6	むすび	5
第3章	GAN	6
3.1	まえがき	6
3.2	Generative Adversarial Networks	6
3.3	Creative Adversarial Networks	6
3.4	Semantically Decomposed GAN	7
3.5	WGAN-GP	8
3.6	むすび	9
第4章	提案手法	10
4.1	まえがき	10
4.2	提案手法	10
4.3	むすび	10
第5章	実験結果	11
5.1	まえがき	11
5.2	使用するデータセット	11
5.3	Creative Adversarial Networks の実行	12
5.4	Semantically Decomposed GAN の実行	13
5.5	WGAN-GP の実行	14
5.6	CNN による評価	15
5.7	考察	16
5.8	むすび	16
第6章	結論と今後の課題	17
6.1	結論	17

6.2 今後の課題	17
謝辞	18
参考文献	19
図一覧	20
表一覧	21
研究業績	22

第1章 序論

1.1 研究の背景

和菓子は五感の芸術であると言われている。これはすべての和菓子の見た目や香り、味、食感がそれぞれ異なり個性を感じられるからである。また同じ植物、同じ風景を題材にした和菓子であっても、使用した材料や製作する職人によってさまざまな表現が存在する。しかし、和菓子において独創的かつ分かり易いデザインを新たに考案することは容易でないと考えられる。

Generative Adversarial Networks (GAN)は2014年に Ian J. Goodfellow によって考案された生成モデルを訓練させるためのモデリング手法である [1]。近年 GAN を用いた Deep Learning による画像生成は様々な分野において行われている。さらに、GAN を改良、発展させることで要求に合った精度の高い画像を生成することができる。したがって GAN を用いることで、和菓子のデザインを考案することができると考えられる。本研究では GAN の発展形である Creative Adversarial Networks (CAN), Semantically Decomposed GAN (SD-GAN) 及び WGAN-GP を用いて和菓子の画像を生成し、Convolutional Neural Network (CNN)を用いて生成画像の評価を行う。

1.2 本研究の目的

和菓子のデザインを GAN によって考案することでデザインのマンネリ化を防ぎ、人間には思いつかない新たなデザインの和菓子も生成できるようになる。GAN は画像データセットを学習することで一度に大量の画像の生成が可能であるが、生成画像が和菓子として成立するかどうかは生成画像を人間の目で見て確認するのは困難である。また主観的な意見が反映されてしまい正確に評価できない可能性がある。CNN を用いて自動で評価することによって定量的な評価を行うことができ、この評価を用いることで精度の高い和菓子画像を生成することが可能である。

1.3 本論文の構成

本論文の構成を以下に示す。

第1章は本章であり、本研究の背景、目的、関連研究について述べる。

第2章ではニューラルネットワーク、特に Convolutional Neural Network (CNN)について述べる。また本研究で作成した和菓子画像を評価する CNN に用いた Fine Tuning と、Fine Tuning に用いたモデル VGG16 について述べる。

第3章では Generative Adversarial Networks (GAN)の概要や仕組みについて述べる。また本研究で用いた Creative Adversarial Networks, Semantically Decomposed GAN

(SD-GAN), WGAN-GP といった 3 種類の GAN の概要, 仕組みについて述べる.

第 4 章では GAN で生成した和菓子の画像の評価方法を提案する.

第 5 章では第 3 章で述べた 3 種類の GAN を和菓子画像のデータセットに対して実行した結果と, 生成された画像を評価するための CNN の作成と, その CNN による評価について述べる. また実験結果について考察を述べる.

第 6 章では本論文の結論と今後の課題について述べる.

第2章 CNN

2.1 まえがき

本章では、本研究の基盤技術であるニューラルネットワークのうち、特に Convolutional Neural Network (CNN)について述べる。また本研究で作成した和菓子画像を評価する CNN に用いた Fine Tuning と、Fine Tuning に用いたモデル VGG16 について述べる。

2.2 ニューラルネットワーク

ニューラルネットワークは生物の脳の神経ネットワークをモデルとしたコンピュータ処理の仕組みであり、入力層、中間層、出力層から構成される [2] [3]。ニューラルネットワークの構造例を図 2.1 に示す。中間層は複数ある場合もある。各層は複数のノードを持ち、2層間のノードはエッジで結ばれる。各ノードは値を持ち、各エッジは重みを持つ。あるノードの値はそのノードと接続する前の層のノードの値、接続エッジの重みの値、その層が持つ活性化関数から計算される。機械学習では、学習データによって最適化されたニューラルネットワークを学習モデルとし、エッジの重みを最適化することで入力値を出力値に変換するように学習する。

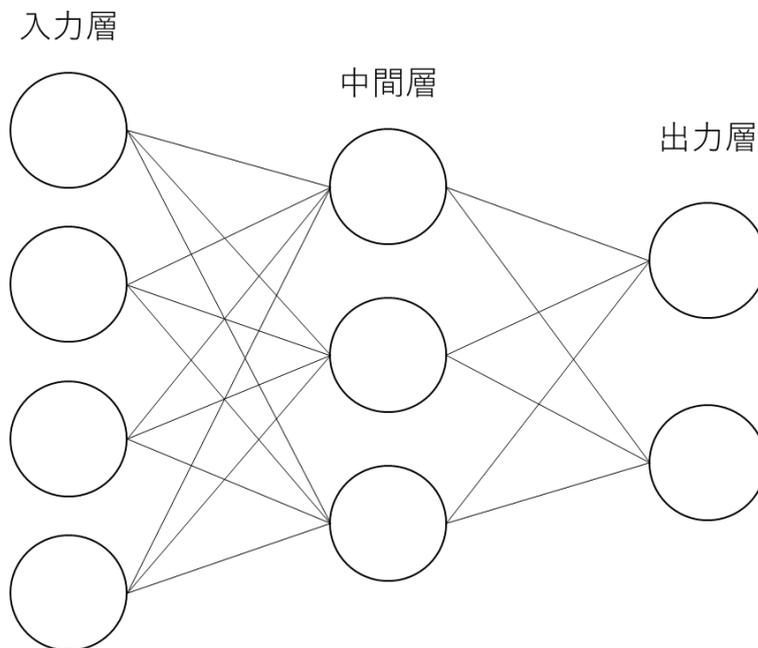


図 2.1 ニューラルネットワークの構造例

2.3 Convolutional Neural Network

Convolutional Neural Network (CNN)は入力層, 畳み込み層, プーリング層, 全結合層, 出力層から構成されるニューラルネットワークである [2] [3] [4] [5]. CNN の構造例を図 2.2 に示す. 畳み込み層, プーリング層, 全結合層は複数あり, 入力層と出力層の間に繰り返される.

畳み込み層は元の画像からフィルタによって特徴抽出を行う層で, 演算に畳み込みを用いる. 畳み込み演算は入力データに対してフィルタのウィンドウを一定間隔でスライドさせて適用する. それぞれのノードにおいてフィルタの要素と入力の対応する要素を乗算して和を求め, 結果を出力に対応するノードに格納する. これをすべてのノードで行う.

プーリング層は入力データを重要な情報を残しつつ縮小する層である. 情報を圧縮することで入力データの微小な位置変化に対して頑健となる. 畳み込み層とプーリング層によって特徴を検出できる.

全結合層は前の層のノードと自分の層のノードがすべてエッジで結ばれている層である. ノードの値は前の層のノードの値, エッジの重み, 活性化関数から計算される.

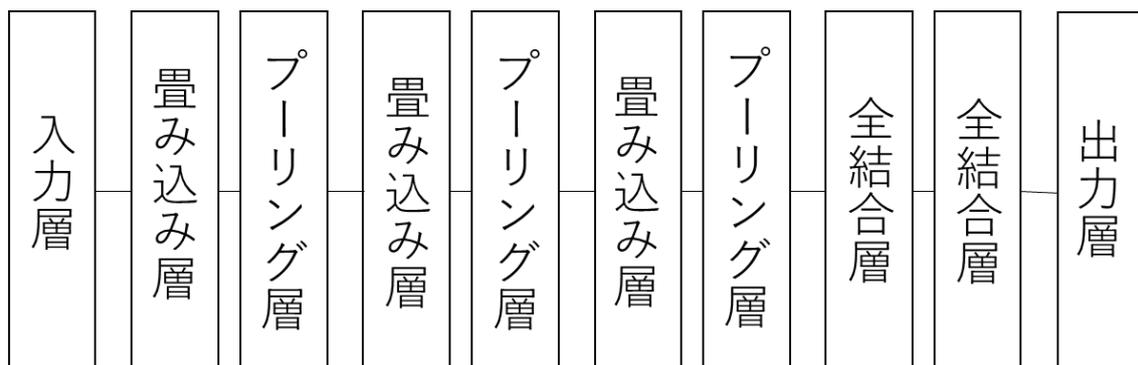


図 2.2 CNN の構造例

2.4 Fine Tuning

Fine Tuning は既に学習済みのモデルの一部を再利用して, 新たなモデルを構築する方法である [6] [7]. Fine Tuning を行うモデルの構造例を図 2.3 に示す. 深層学習では大量の学習データが必要であるが, あらかじめ他の大量のデータによって学習されたモデルのパラメータの重みを新しいモデルのパラメータの初期値として用いることで, 新しいモデルは少ない量の学習データでも構築できる. 異なる学習データであってもモデルを再利用できるのは, 大量の画像から大量のクラスに分類したネットワークの下層部分はデータの内容が異なっても類似するためである.

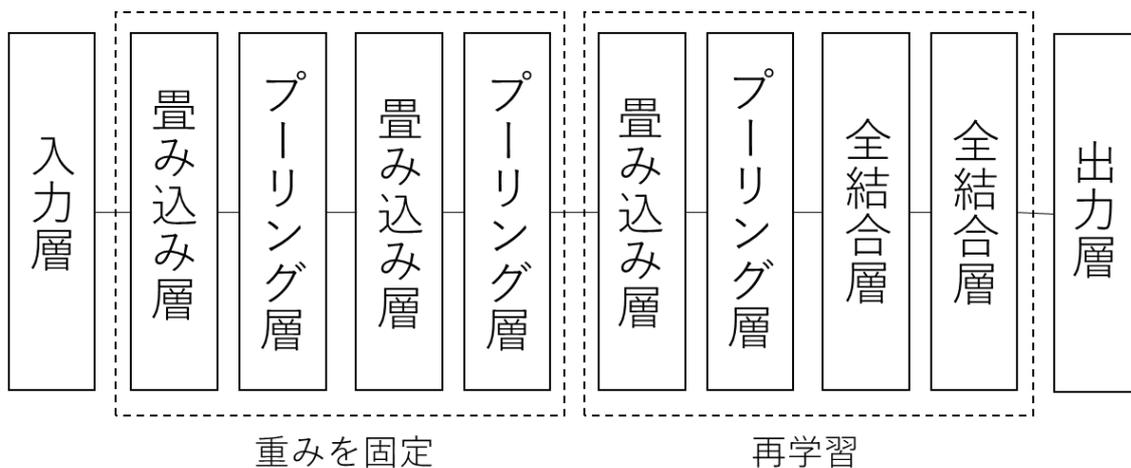


図 2.3 Fine Tuning の例

2.5 VGG16

VGG16 は大規模画像データセット ImageNet によって学習された畳み込み層 13 層と全結合層 3 層の計 16 層からなる CNN モデルである [3] [6]. VGG16 モデルの構造を図 2.4 に示す. 入力された画像を 1000 クラスに分類することができる.



図 2.4 VGG16 モデルの構造

2.6 むすび

本章では, 本研究の基盤技術であるニューラルネットワークのうち, Convolutional Neural Network (CNN), Fine Tuning, VGG16 について述べた.

第3章 GAN

3.1 まえがき

本章では, Generative Adversarial Networks (GAN)の概要や仕組みについて述べる. また本研究で用いた Creative Adversarial Networks (CAN), Semantically Decomposed GAN (SD-GAN), WGAN-GP といった 3 種類の GAN の概要, 仕組みについて述べる.

3.2 Generative Adversarial Networks

Generative Adversarial Networks (GAN)は, 2014 年に Ian J. Goodfellow によって考案された深層学習を用いて生成モデルを訓練させるためのモデリング手法である [1]. GAN の構造を図 3.1 に示す. 学習は Generator と Discriminator と呼ばれる 2 種類のネットワークを敵対的にトレーニングする. Generator は偽物の画像を生成するネットワークであり, 学習用データに近い画像を生成することで Discriminator を欺くように学習する. Discriminator は Generator が生成した画像を学習用データセットと比較し生成画像が本物かどうかを識別するネットワークあり, 偽物の画像と本物の画像を見分けられるように学習する. このように二つのネットワークを競合させることで精度の高い画像を生成できるようになる.

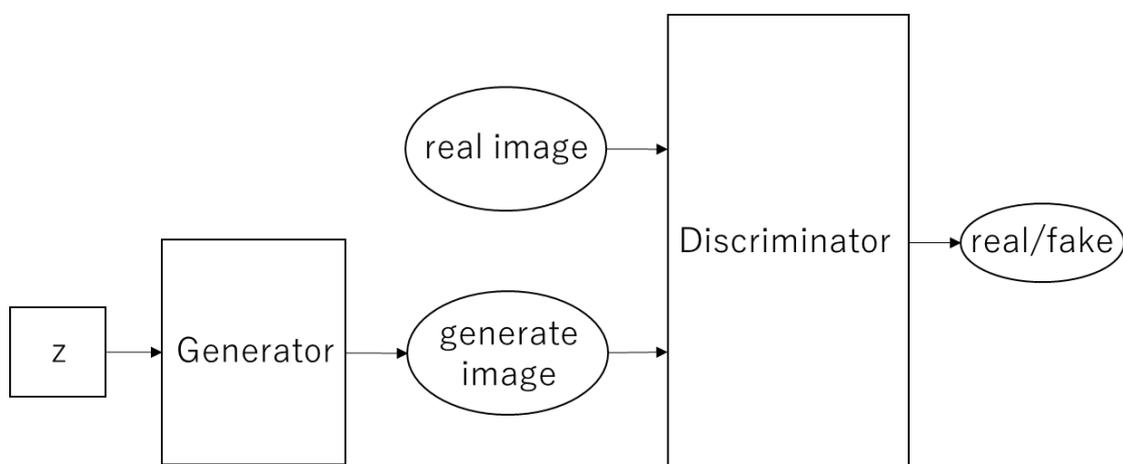


図 3.1 GAN の構造

3.3 Creative Adversarial Networks

Creative Adversarial Networks (CAN)は2017年にラトガーズ大学の Ahmed Elgammal らが発表した GAN を基本とした画像生成技術で, 新しい芸術を生み出すことができる [8]. CAN の構造を図 3.2 に示す. GAN と同様に Generator と Discriminator を敵対的にトレ

ーニングする。Generator が画像を生成し、Discriminator は絵画のスタイルごとに分類された画像データセットと生成画像を比較して芸術らしさと学習したスタイルへの適応度を学習する。芸術らしさを持つ画像かつ学習したスタイルへの適応度が低い画像を正解とすることによって、既存のスタイルから逸脱した芸術作品の画像の生成を行う。

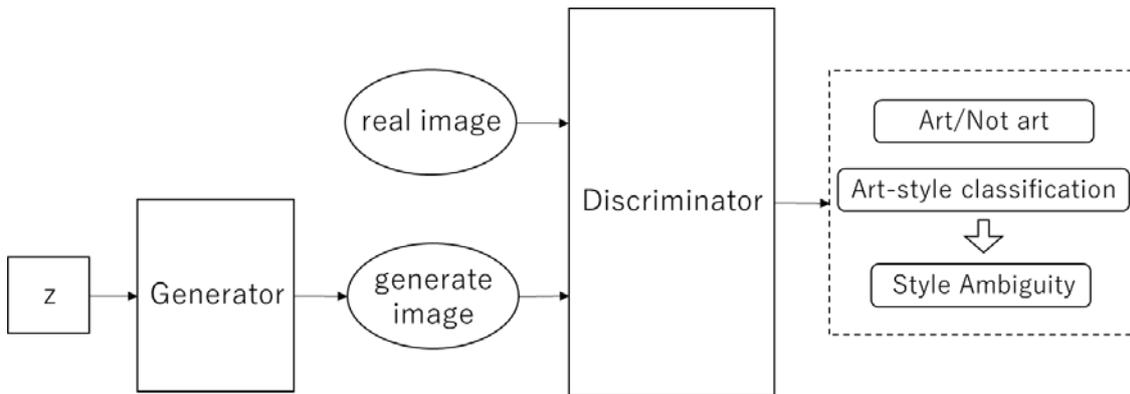


図 3.2 CAN の構造

3.4 Semantically Decomposed GAN

Semantically Decomposed GAN (SD-GAN)は 2018 年にカリフォルニア大学の Chris Donahue らが発表した GAN を基本とした画像生成技術である [9]。SD-GAN の構造を図 3.3 に示す。一つの被写体のライティング、ポーズなどの異なる複数の画像を生成することができ、新しい被写体の画像を生成することもできる。GAN と同様に Generator と Discriminator を敵対的にトレーニングする。SD-GAN が GAN と異なる点は 2 枚の画像の被写体が同一であるかどうかの判別を行うために Siamese Network を用いている点である。Siamese Network は二つの同一のニューラルネットワークを用いて入力データの類似度を比較するネットワークである。Siamese Network を用いて Generator と Discriminator を二つずつ使用することで特徴が異なっても同一被写体かどうかの判別を行っている。

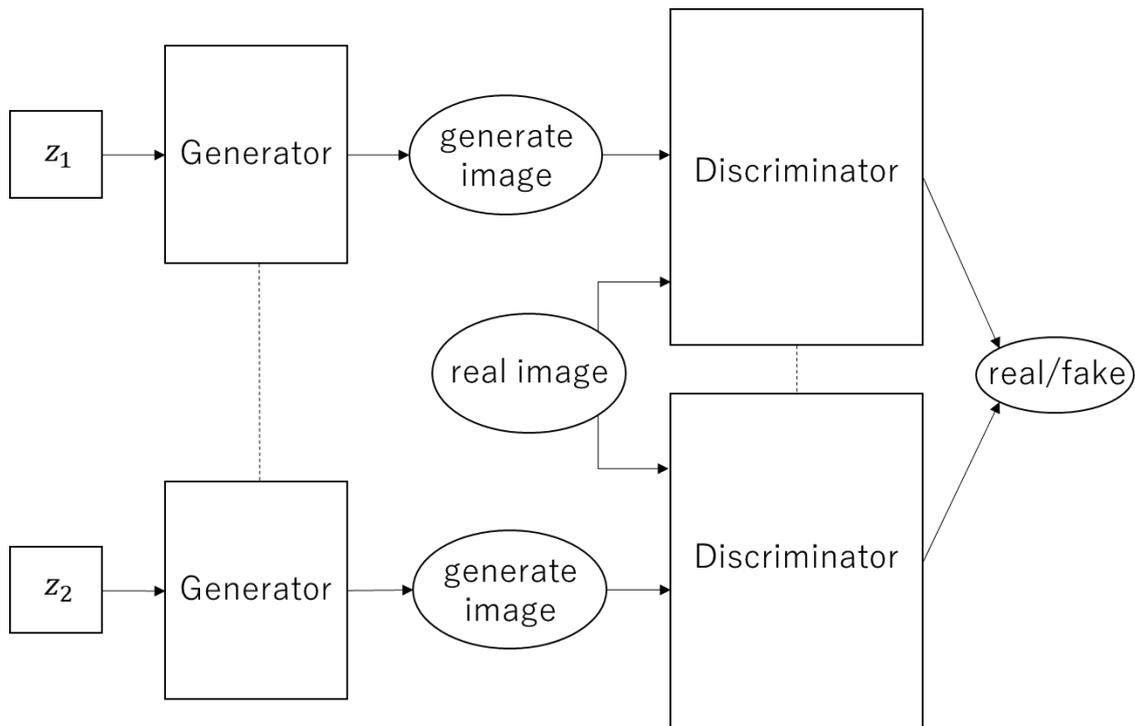


図 3.3 SD-GAN の構造

3.5 WGAN-GP

2017年にクーラント数理科学研究所のMartin Arjovskyらによって Wasserstein GAN (WGAN)が提案された [11]. GANと同様に Generator と Discriminator を敵対的にトレーニングする.

GANは Discriminator の学習が上手く進まない場合に Generator が正確に学習できなかつたり, Discriminator の学習が速く損失関数の勾配が消失して Generator の学習が進まなくなつたり, 同じ出力を生成するようになってしまうモード崩壊が生じたりと学習が不安定になる. GAN と WGAN の異なる点は GAN の Discriminator は画像が本物であるか偽物であるかを識別できるように学習するが, WGAN は Wasserstein 距離を計算できるように学習する点である. これにより WGAN は GAN よりも安定した学習を実現できる. しかし WGAN はリップシツク制約をみたすために weight clipping が行われるが, clipping parameter が大きいと Discriminator の学習が上手く進まなかつたり, clipping parameter が小さいと勾配が消失したりと学習が依然として不安定であった.

そこで WGAN を改善した WGAN-GP が 2017年にモントリオール学習アルゴリズム研究所の Ishaan Gulrajani らによって提案された [10]. WGAN-GP はリップシツク制約をみたすために weight clipping の代わりに gradient penalty を使用する. gradient penalty は損失関数に勾配のノルムが 1 以外のときにペナルティを与えるというものである. これに

より安定した学習を行うことができる。

3.6 むすび

本章では Generative Adversarial Networks (GAN)の概要や仕組みについて、また本研究中で用いた Creative Adversarial Networks(CAN), Semantically Decomposed GAN (SD-GAN), WGAN-GP といった 3 種類の GAN の概要, 仕組みについて述べた。

第4章 提案手法

4.1 まえがき

本章では，CNN を用いて GAN で生成した和菓子の画像を評価する方法を提案する．

4.2 提案手法

GAN で和菓子画像のデータセットを用いて生成した画像が和菓子とみなせるかを，CNN を用いて評価する．まず，CNN に和菓子の画像と和菓子でない物体の画像を学習させる．識別精度の高い CNN を構築するために，学習には VGG16 モデルを用いて Fine Tuning を行う．Fine Tuning を行った VGG16 モデルの構造を図 4.1 に示す．全結合層とその一つ前の畳み込み層のブロックの重みを再学習し，それ以前の層の重みは固定する．

学習終了後，GAN によって生成された画像を学習した CNN に入力し和菓子画像であるかどうかを判別する．複数枚生成画像を入力して和菓子画像と認識された画像の割合を求めることによって正解率を算出する．

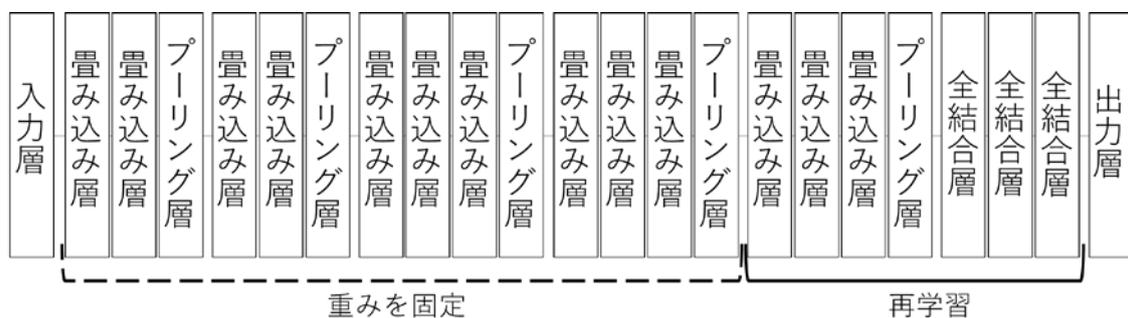


図 4.1 VGG16 の Fine Tuning

4.3 むすび

本章では GAN による生成画像の評価方法についての提案手法を述べた．

第5章 実験結果

5.1 まえがき

本章では, 3種類のGANを和菓子画像のデータセットを用いて実行した結果と生成された画像を評価するCNNを作成とそのCNNによる評価について述べる. また, 実験結果について考察を述べる.

5.2 使用するデータセット

本研究で使用する画像データセットは, インターネット上で和菓子の画像を取得し作成した. 27種類の素材の画像を計11051枚収集した. データセットに使用した画像の素材ごとの枚数を表5.1に示す.

表 5.1 データセットに使用した画像

素材	画像枚数[枚]
朝顔	589
紫陽花	913
うぐいす	198
うさぎ	381
団扇	494
梅	308
カーネーション	406
桔梗	252
菊	770
金魚	744
栗	486
コスモス	284
桜	658
山茶花	270
菖蒲	456
水仙	359
つつじ	211
椿	180
手鞠	294
撫子	231
花火	187
薔薇	527
向日葵	535
藤の花	276
牡丹	216
紅葉	452
桃	373
合計	11051

5.3 Creative Adversarial Networks の実行

和菓子のデータセットを用いて CAN を実行する。CAN はスタイルごとに分類されたデータセットを使用するため、和菓子の素材をスタイルとみなし 27 種類のスタイルに分類し

た. 学習時の画像サイズは 64[pixel], ミニバッチサイズは 16, エポック数は 100 であった. CAN による生成画像例を図 5.1 に示す.

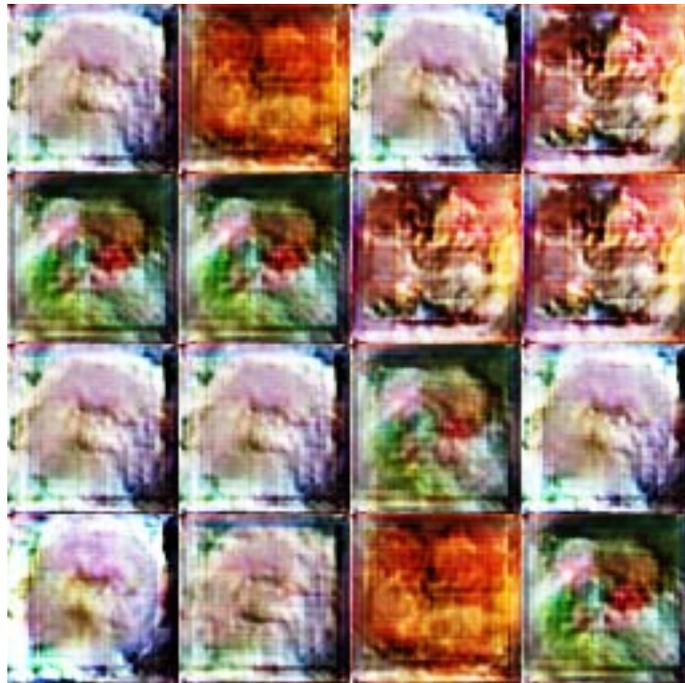


図 5.1 CAN による生成画像例

5.4 Semantically Decomposed GAN の実行

和菓子データのデータセットを用いて SD-GAN を実行する. データセットのうち学習用に 8841 枚, 検証用に 1105 枚, テスト用に 1105 枚使用した. 学習時の画像サイズは 64[pixel], ミニバッチサイズは 16, エポック数は 100 であった. SD-GAN による生成画像例を図 5.2 に示す.

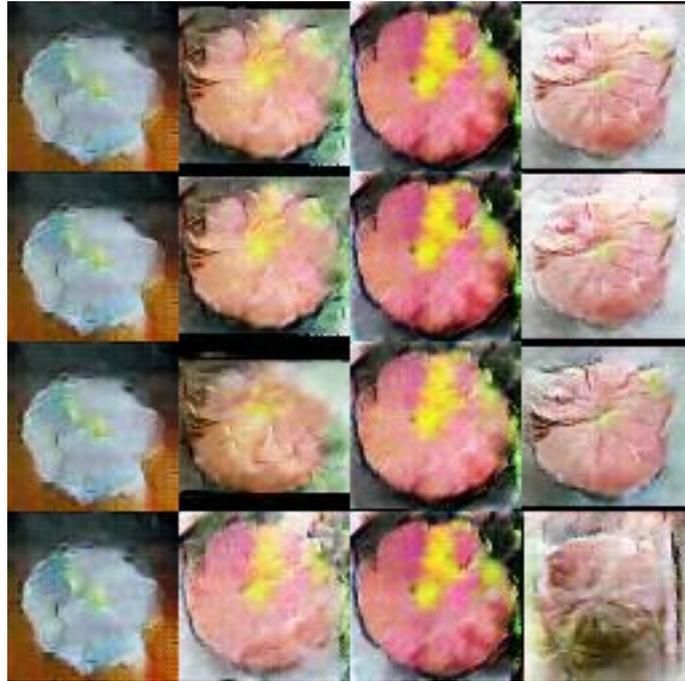


図 5.2 SD-GAN による生成画像例

5.5 WGAN-GP の実行

和菓子のデータセットを用いてWGAN-GPを実行する. 学習時の画像サイズは64[pixel], ミニバッチサイズは16, エポック数は100であった. WGAN-GPによる生成画像例を図5.3に示す.



図 5.3 WGAN-GP による生成画像例

5.6 CNN による評価

CNN を用いて入力された画像が和菓子画像か否かを判定する．CNN の学習は VGG16 モデルを用いて Fine Tuning を行った．CNN の学習に使用した画像は，5.2 で述べた和菓子画像データセットの中から無作為に選択した 11000 枚と，和菓子の画像でない物体の画像 11000 枚の計 22000 枚である．このうち，和菓子画像 8800 枚と負例画像 8800 枚を学習用画像，和菓子画像 2200 枚と負例画像 2200 枚を評価用画像とした．学習時のエポック数は 50 であった．学習後，評価用画像による正解率は 98.91[%]であった．

CAN, SD-GAN, WGAN-GP によって生成されたそれぞれ 1000 枚の画像を CNN によって評価し，正解率を算出した．CAN による生成画像の正解率は 7.9[%]，SD-GAN による生成画像の正解率は 81.9[%]，WGAN-GP による生成画像の正解率は 80.5[%]であった．正解率を表 5.2 に示す．

表 5.2 生成画像の評価結果

	正解率[%]
CAN	7.9
SD-GAN	81.9
WGAN-GP	80.5

5.7 考察

CNN の学習は, 評価用画像による正解率が 98.91[%]であったことから成功したといえる. また, GAN による生成画像の評価結果は CAN による生成画像の正解率が低く, SD-GAN による生成画像と WGAN-GP による生成画像の正解率は高く, 一番正解率が高いのは SD-GAN による生成画像であった. CNN による評価の結果からは, 和菓子画像の生成に最も適しているのは SD-GAN であるといえる.

一方で各 GAN の生成画像例を比較すると, CAN による生成画像と SD-GAN による生成画像は類似した画像が多く生成されている. CAN においてはモード崩壊が生じているため類似した画像が生成される. SD-GAN においては構造上一つの被写体のポーズやライティングの異なる画像を生成するため類似した画像が生成される. WGAN-GP による生成画像は, 類似した画像が少なく多様な画像が生成できた. 目視による評価の結果からは, 和菓子画像の生成に最も適しているのは WGAN-GP であるといえる.

CNN による評価は生成画像一枚ずつが和菓子であるかどうかを判別しているため, 生成画像の類似性についての評価ができない. したがって目視による評価と異なる結果になった. CNN による評価において SD-GAN による生成画像の正解率は 81.9[%], WGAN-GP による生成画像の正解率は 80.5[%]であり, WGAN-GP による生成画像も十分に高い正解率であるといえる. したがって和菓子画像の生成に最も適しているのは WGAN-GP であるといえる.

5.8 むすび

本章では, 3種類の GAN を和菓子画像のデータセットを用いて実行した結果と生成された画像を評価する CNN を作成とその CNN による評価について述べた. また, 実験結果について考察を述べた. 実験結果より, 和菓子画像の生成に最も適しているのは WGAN-GP であることが分かった.

第6章 結論と今後の課題

6.1 結論

本研究では3種類のGANによって和菓子画像を生成し,作成した和菓子画像判別のCNNを用いて生成画像の正解率を評価した. CNNの学習はVGG16によるFine Tuningを行うことで正解率98.91[%]のCNNモデルを生成することができた.

また,和菓子画像判別CNNを用いて算出したGANによる生成画像の正解率はCANによる生成画像は7.9[%],SD-GANによる生成画像は81.9[%],WGAN-GPによる生成画像は80.5[%]であった.しかし,CAN及びSD-GANによる生成画像は類似した画像が多く,多様性が低い.WGAN-GPによる生成画像は多様性が高く,CNNによる正解率も十分に高いため和菓子画像の生成に一番適しているのはWGAN-GPであることが分かった.

6.2 今後の課題

本研究で作成したCNNモデルは入力画像1枚に対して和菓子であるか否かを判別するモデルであるため生成画像全体の多様性に対しての評価ができなかった.今回は目視により多様性に対しての評価を行ったが,複数枚の画像の多様性に対して定量的な評価をする方法を検討すべきであると考えられる.

謝辞

本研究の実験環境を整えてくださり、研究の方向性を示して下さったり問題点を提起して下さったりと適切な指導を賜った渡辺裕教授に心より感謝申し上げます。

また、日頃からご意見やアドバイスをくださった研究室の皆様に御礼申し上げます。

最後に、私をここまで育ててくれ、支えてくれた家族に感謝します。

参考文献

- [1] I. J. Goodfellow, J. P. Abadie, M. Mirza, B. Xu, D. W. Farley, S. Ozair, A. Courville and Y. Bengio, "Generative Adversarial Nets," Conference and Workshop on Neural Information Processing Systems (NIPS), pp.1-9, Dec. 2014.
- [2] 武井宏将, 初めてのディープラーニング—オープンソース"Caffe"による演習付き, pp.76-78, 81-86, リックテレコム, 2016.
- [3] 斎藤康毅, ゼロから作る Deep Learning—Python で学ぶディープラーニングの理論と実装, pp.39, 205-210, 219-221, 250, オライリー・ジャパン, 2016.
- [4] “定番の Convolutional Neural Network をゼロから理解する,” DeepAge, 2016 年 11 月 7 日. https://deepage.net/deep_learning/2016/11/07/convolutional_neural_network.html. 2019 年 1 月閲覧.
- [5] システムインテグレータ, “畳み込みニューラルネットワーク__CNN(Vol.16),” AISIA, 2018 年 5 月 11 日. <https://products.sint.co.jp/aisia/blog/vol1-16#toc-4>. 2019 年 1 月閲覧.
- [6] “少ない画像から画像分類を学習させる方法 (keras で転移学習: fine tuning),” SPJ, 2018 年 9 月 26 日. <https://spjai.com/keras-fine-tuning/>. 2019 年 1 月閲覧.
- [7] S. Yuki, “Keras で学ぶ転移学習,” Elix Tech Blog, 2016 年 6 月 22 日. <https://elix-tech.github.io/ja/2016/06/22/transfer-learning-ja.html>. 2019 年 1 月閲覧.
- [8] A. Elgammal, B. Liu, M. Elhoseiny and M. Mazzone, "CAN: Creative Adversarial Networks Generating "Art" by Learning About Styles and Deviating from Style Norms," ICCV, pp.1-22, June 2017.
- [9] C. Donahue, Z. C. Lipton, A. Balsubramani and J. McAuley, "Semantically Decomposing the Latent Spaces of Generative Adversarial Networks," arXiv preprint arXiv:1705.07904, 2017.
- [10] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. Courville, "Improved Training of Wasserstein GANs," arXiv preprint arXiv: 1704.00028, 2017.
- [11] M. Arjovsky, S. Chintala and L. Bottou, "Wasserstein GAN," arXiv preprint arXiv:1701.07875, 2017.

図一覧

図 2.1	ニューラルネットワークの構造例	3
図 2.2	CNN の構造例	4
図 2.3	Fine Tuning の例	5
図 2.4	VGG16 モデルの構造	5
図 3.1	GAN の構造	6
図 3.2	CAN の構造	7
図 3.3	SD-GAN の構造	8
図 4.1	VGG16 の Fine Tuning	10
図 5.1	CAN による生成画像例	13
図 5.2	SD-GAN による生成画像例	14
図 5.3	WGAN-GP による生成画像例	15

表一覧

表 5.1	データセットに使用した画像.....	12
表 5.2	生成画像の評価結果.....	15

研究業績

1. 大山優香, 渡辺裕, “GAN による和菓子画像の生成と評価”, 電子情報通信学会総合大会, 2019 年 3 月発表予定