敵対的生成ネットワークを用いた和菓子画像生成の検討
A Study on Generation of Japanese Sweets Images by Generative Adversarial Networks

大山 優香† 渡辺 裕†
Yuka Oyama† Hiroshi Watanabe†

早稲田大学大学院 基幹理工学研究科 情報理工・情報通信専攻†
Department of Computer Science and Communications Engineering, Waseda University†

Abstract: 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はないと考えられる．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．GAN は Generator と Discriminator を競争させながら互いに学習させることで訓練データに近いデータを生成できるようになる．近年 GAN を用いた Deep Learning による画像生成はさまざま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．そこで本研究では，GAN を用いた和菓子の画像生成について検討する．

1 はじめに

和菓子は植物や動物，風景を題材にしたものが多く，その和菓子から四季を感じることができる．和菓子のデザインは同じ題材の和菓子であっても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，新たに和菓子のデザインを考案することは容易ではない．そこで，和菓子のデザインを Generative Adversarial Networks（GAN)を用い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 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]．GAN は画像データセットを学習することで Generator と Discriminator を競争させながら互いに学習し，訓練データに近いデータを生成できるようになる．そこで，本研究では Semantically Decomposed GAN(SD-GAN)を用いて和菓子の画像を生成する手法について検討する．

2 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 をベースとした画像生成技術である[2]．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．

3 MS-SSIM

Structural similarity(SSIM)は画像評価指標の一つであり，2 つの画像間の類似性を測定するために使用される[3]．SSIM は画像内の小領域ごとに算出される．SSIM を画像全体で算出するものが multi-scale structural similarity(MS-SSIM)である[3]．画像の全体的な品質を測定することができる．

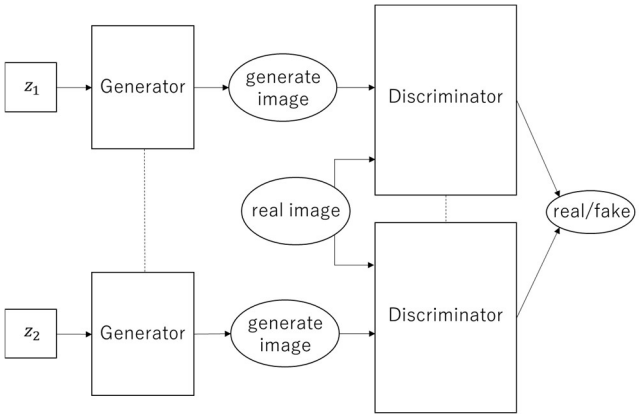

図 1: Semantically Decomposed GAN

4　提案手法

本研究では，和菓子の画像データセットを用いて SD-GAN を学習し，和菓子画像を生成する．2 種類の方法でデータセットを分類することで，SD-GAN を使用して和菓子画像を生成するのに適切な分類方法を検討する．

5　実験

5.1 実験概要

和菓子画像 11052 枚のデータセットを用いて SD-GAN を学習させる．使用したデータセットは 27 種類の題材の和菓子画像のデータセットである．2 種類の方法でデータセットの分類を行う．方法 1 の分類方法は題材ごとに 27 種類に分類し，1 グループあたりの平均画像枚数は約 400 枚である．方法 2 の分類方法は一つの題材内で材料や形状により細分化し，1 グループあたりの平均画像枚数は約 8 枚である．

また，CNN と MS-SSIM を用いて生成画像の評価を行う．生成画像 1000 枚を，和菓子画像を学習した CNN によって評価し正解率を算出する．生成画像 1000 枚から無作為に選択した 2 枚の画像対 100 対の画像間の MS-SSIM を測定し，平均した値によって多様性を評価する．

5.2 実験結果

方法 1 のデータセットの分類による SD-GAN の生成画像例を図 2 に示す．方法 2 のデータセットの分類による SD-GAN の生成画像例を図 3 に示す．CNN による生成画像の和菓子判別率を表 1 に示す．MS-SSIM による生成画像の多様性評価結果を表 2 に示す．CNN による和菓子判別率は方法 2 が 14.4 ポイント高いが，逆に MS-SSIM の値は 0.04 ポイント悪化した．

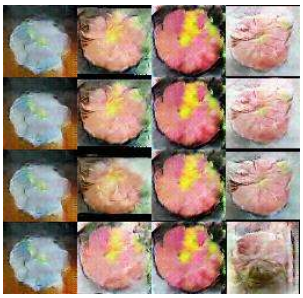

図 2: 方法 1

図 3: 方法 2

表 1: 和菓子判別率

	正解率[%]
方法 1	81.9
方法 2	96.3

表 2: MS-SSIM

	MS-SSIM
方法 1	0.36
方法 2	0.40

6　まとめ

本稿では SD-GAN を用いた和菓子画像の生成を行った．データセットを細分化することで，和菓子判別率は上昇したが多様性評価である MS-SSIM は悪化した．和菓子画像生成に最適な GAN の検討を行いたい．

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] 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.

[3] Zhou Wang, Alan C. Fellow, Hamid R. Sheikh, Eero P. Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE, pp.600-612, 2004

早稲田大学 基幹理工学部 情報通信学科 渡辺研究室
〒169–0072 東京都新宿区大久保 3-14-9 早大シルマンホール 401
Phone: 03–5286–2509
E-mail: yuuchan-ooyamat@ruri.waseda.jp