



# 大量のTwitter上の画像を用いた Conditional Cycle GAN を用いた食事画像カテゴリ変換

堀田 大地, 丹野良介, 下田和, 柳井啓司

## 背景

近年, 深層学習を用いた画像生成・変換に関する研究が熱い, しかし...

文字, 顔, 居住がメイン  
ファッション画像への応用も

2014 2015 2016 2017

食事画像研究にも応用したい...!!

### 目的: 敵対的生成ネットワーク(GAN)による食事画像変換

Image-to-Image変換手法検討

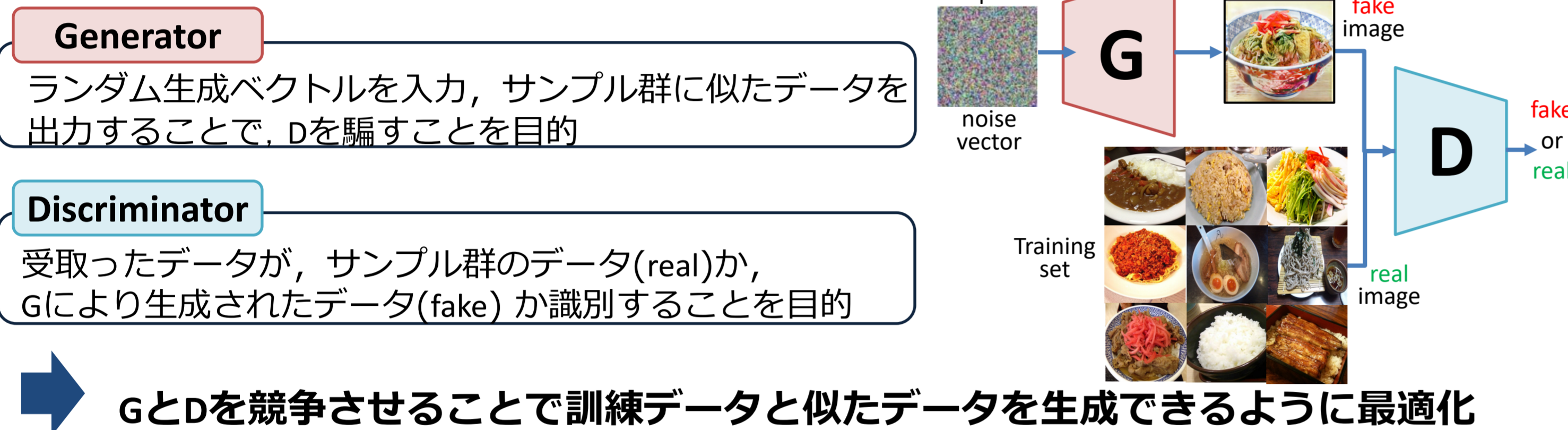
- Generative Adversarial Net
- Neural Style Transfer

Domain Select  
One-hot Vector signal

## デモ動画 1

## 手法 1 : conditional CycleGAN

Generative Adversarial Network(GAN)[1]: 深層学習+生成モデル  
用意したサンプルデータ群に似たデータを生成



Related work ① : pix2pix [CVPR2017]

### GANを用いたImage-to-Image変換手法

手法: Conditional Adversarial Net

変換前と変換後のペア画像を学習をcGANの学習データに用いる

解決: ペア画像さえあれば変換可能

課題: ペア画像が必須

Related work ② : CycleGAN[1] [ICCV2017]

### ドメインxからドメインyへ画像を変換する Image-to-Image変換手法

手法: Cycle Consistency Loss to push  $F(G(X)) \approx X$  (and vice versa)

[元の画像]と, その画像を[変換 -> 逆変換]してもとに戻したものの間の誤差で学習

解決: ペア画像無くても学習可能に

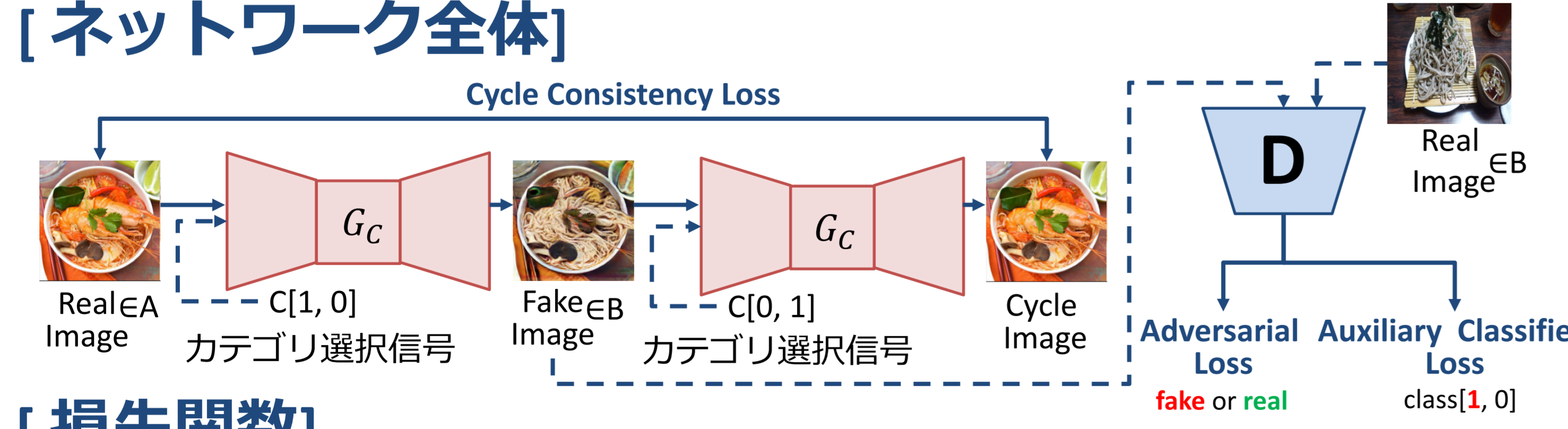
課題: 1:1変換しか出来ない

Related work ③ : ACGAN[2] [ICML2017]

### 識別器(D)に識別タスクを付与することで画像クオリティの向上

手法: Auxiliary Classifier Loss to make high quality image

識別器(D)に真偽判定(Adversarial Loss)だけでなく, クラスの識別をさせる補助的なタスクを追加



### [損失関数]

識別器(D)の損失関数:  $L_D = L_{adversarial} + \lambda_{classifier} L_{classifier}$

変換器(G)の損失関数:  $L_G = L_{adversarial} + \lambda_{classifier} L_{classifier} + \lambda_{cycle} L_{cycle}$

損失重み:  $\lambda_{classifier}, \lambda_{cycle}$

## まとめと今後の課題

- CycleGANの拡張による1対多変換器(cCycleGAN)の作成
  - 今後の課題
  - より多くの手法との比較比較と定量的比較
  - 「井」制約の緩和
  - 任意カテゴリ(学習済みカテゴリ以外)への拡張
- 

## 手法 2 : Multi Style Transfer[3]

Neural Style Transfer[Gatys+ CVPR2016]

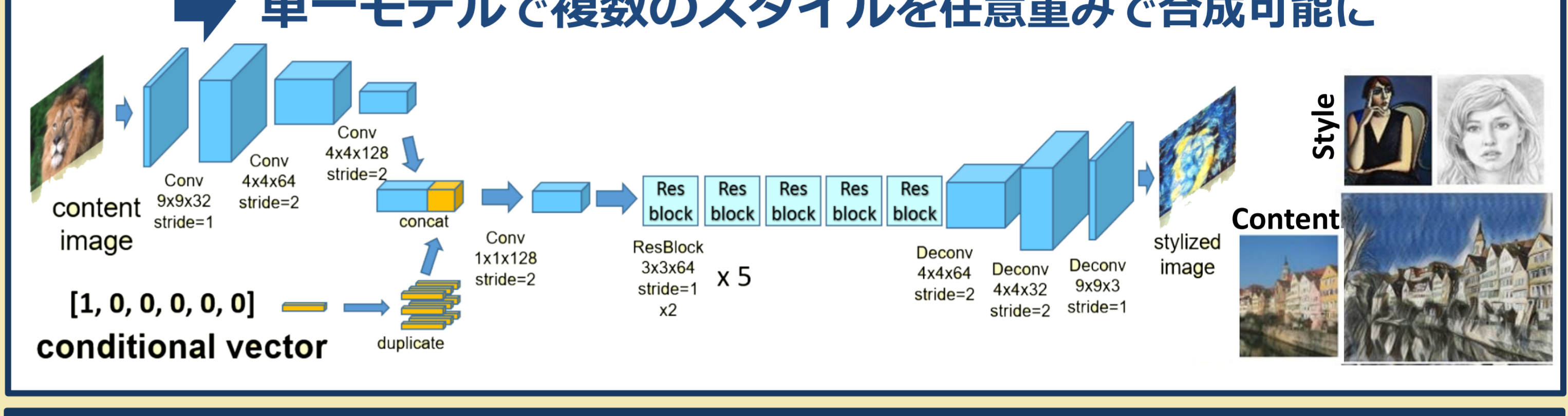
形状を保持したまま画風を転写  
順伝搬及び誤差逆伝播を繰り返す  
GPUで画像の生成に数十秒程度

Content  $F^c$ , Style  $\{G^s | l_{e_s}\}$

Fast Style Transfer[Johnson+ ECCV2016]

特定の画風変換を準伝搬で行う CNN を学習  
非常に高速に画像を変換可能  
1つのモデルで単一の画風変換のみ学習  
消費メモリの増大, 学習に時間要

Style Target  $F^s$ , Content Target  $F^c$ , Image Transform Net



## 実験 : conditional CycleGAN vs Multi Style Transfer[3]

### [データセット]

学習データに多様性が有りすぎると学習が難しい...

|         |         |         |         |         |
|---------|---------|---------|---------|---------|
| カレー     | 炒飯      | 牛丼      | 冷やし中華   | ミートスパ   |
| 34,216枚 | 27,854枚 | 18,396枚 | 13,499枚 | 7,138枚  |
| ラーメン    | 白飯      | 蕎麦      | うなぎ重    | 焼きそば    |
| 74,007枚 | 24,760枚 | 3,350枚  | 5,329枚  | 24,760枚 |

「井」という制約を付けてデータを収集(From Twitter)

ノイズ除去

認識率が高い順にランキング

keyword = '牛丼'で収集

牛丼 (93.7%)

Food 10 categories. Total 230,053 images.

形状が似た画像群が得られる

### [結果] conditional CycleGAN (⊙食事領域のみ別のカテゴリへと変換できた)



### [結果] Multi Style Transfer (⊗食事領域以外も変わってしまった)



## 実験 : 画像枚数の違いによるクオリティへの影響結果

### [結果] ⊙ 画像枚数が多いほうが綺麗に変換ができた



[1] J. Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proc. of ICCV2017.  
 [2] A. Odena, C. Olah, and J. Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. In Proc. of ICML2017.  
 [3] R. Tanno and K. Yanai, DeepStyleCam: A Real-time Style Transfer App on iOS. In Proc. of The 24th International Conference on Multimedia Modeling, 2017.