

# Food Category Transfer with Conditional Cycle GAN and a Large-scale Food Image Dataset

Daichi Horita Ryosuke Tanno Wataru Shimoda Keiji Yanai  
The University of Electro-Communications,  
Tokyo, Japan



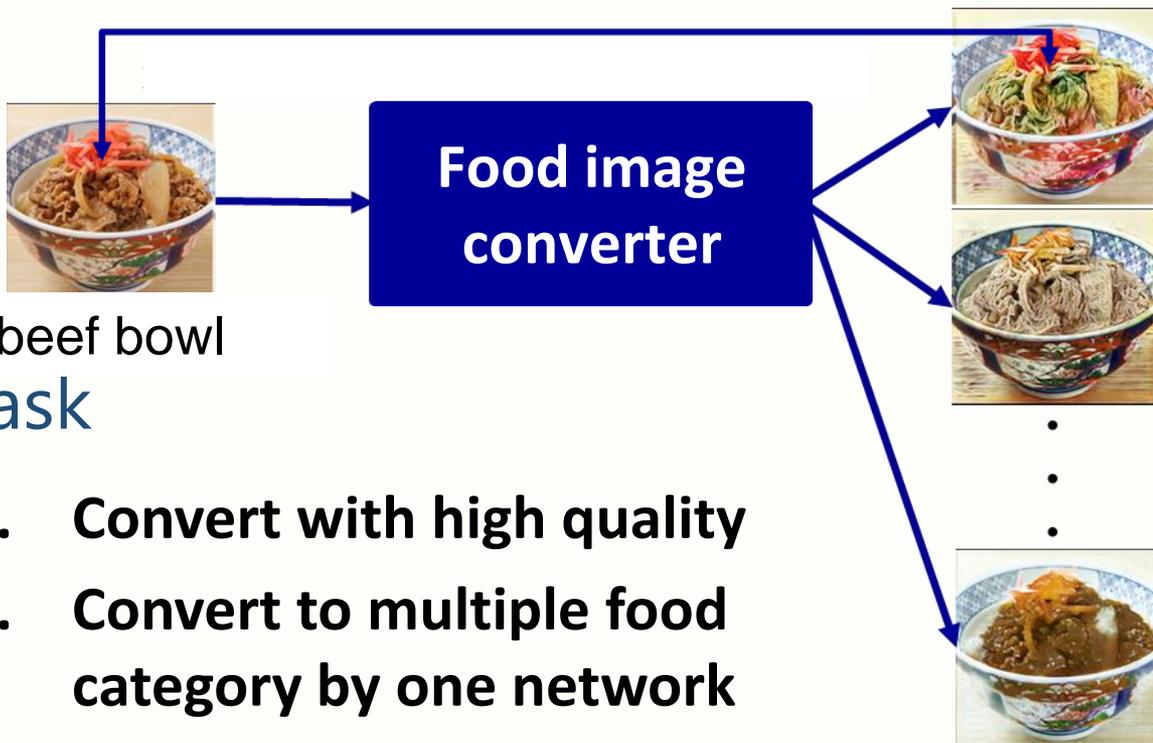
# Image generation method : GAN

- GAN (Generative Adversarial Network)  
[Goodfellow et al. NIPS 2014]
- DCGAN (Deep Convolutional GAN) [Radford et al. ICLR 2016]



# Objective

- Keep shapes of food, which is before conversion



beef bowl

- Task

1. Convert with high quality
2. Convert to multiple food category by one network

Chilled noodles  
Like beef bowl  
干丹

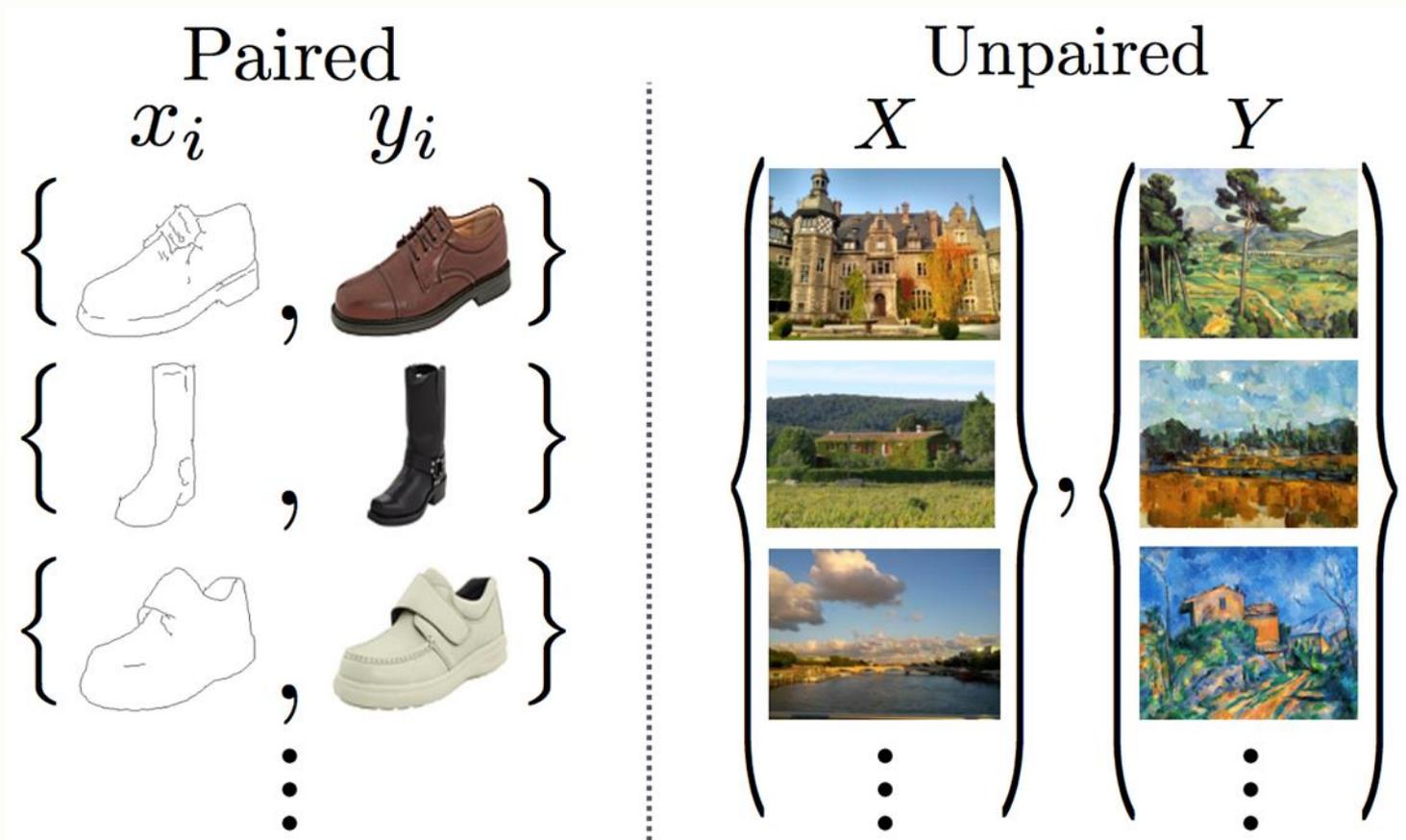
soba  
Like beef bowl  
.  
.  
.

cary  
Like beef bowl

**➡ Domain transfer(CycleGAN)**

# CycleGAN

- CycleGAN([Zhu+ ICCV-17])
  - Cycle GAN is trained with unpaired image set

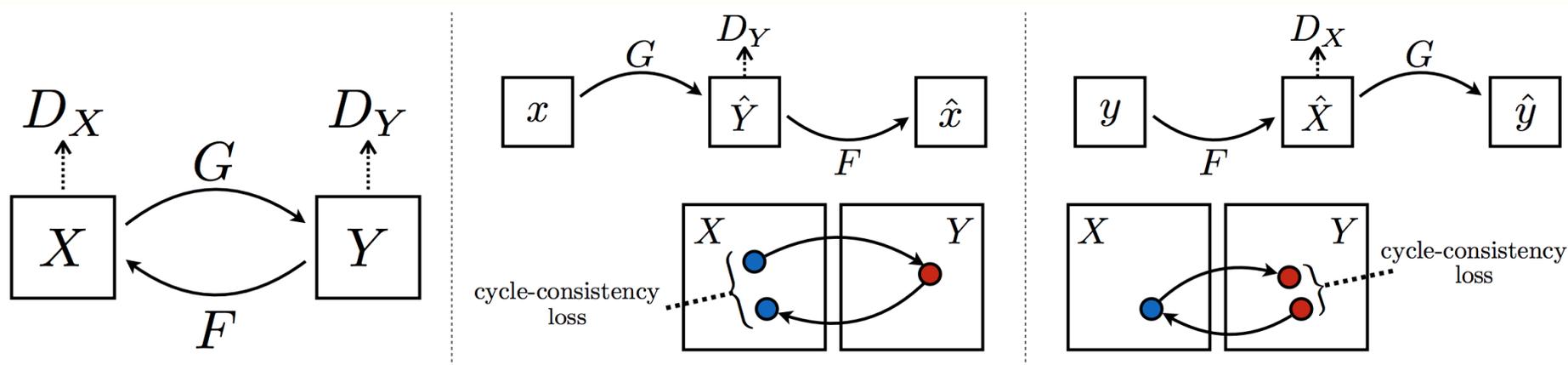


# CycleGAN

- Convert input image domain form  $X$  to  $Y$  by network  $G$ .
- Reverse the domain form  $Y$  to  $X$  by network  $F$ .
- minimize the difference between

the input image  $x$  and reversed image  $\hat{x}$

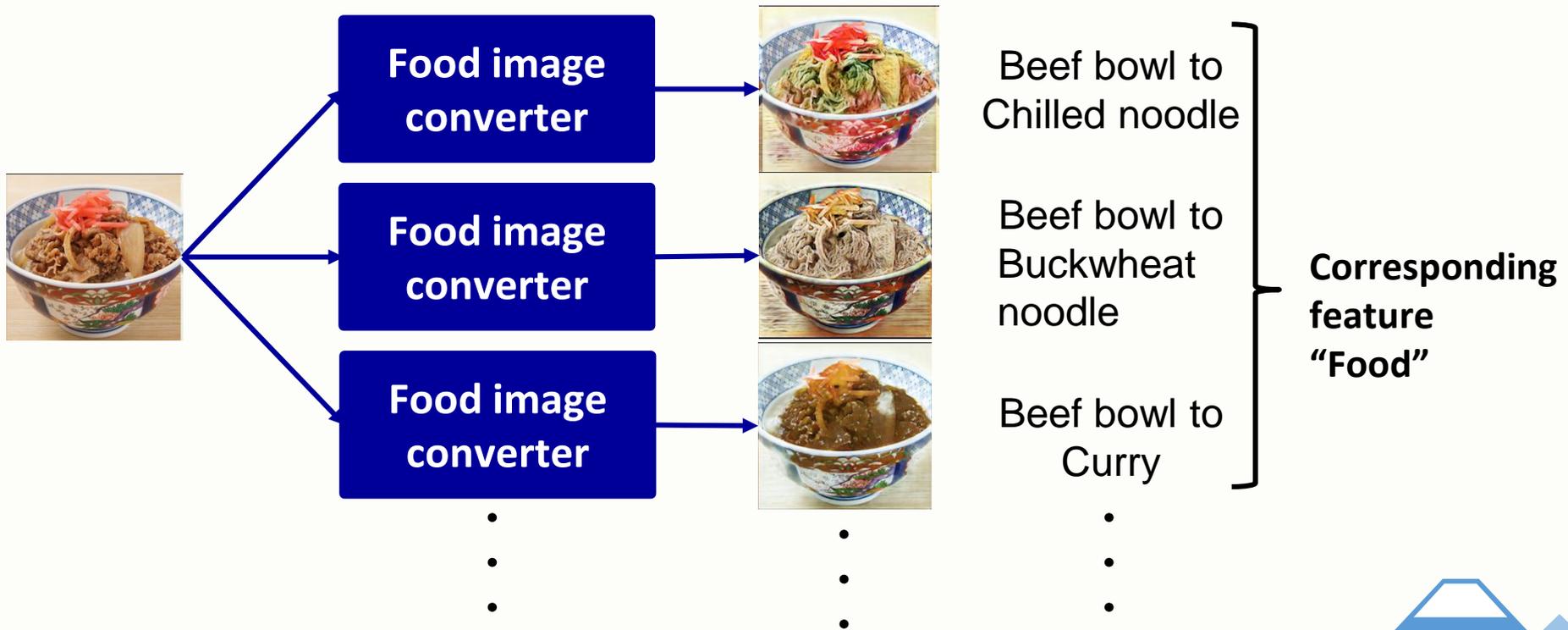
**➔ Conversion can be learned without pair images**  
**Cycle Consistency Loss**



# Limitation of CycleGAN

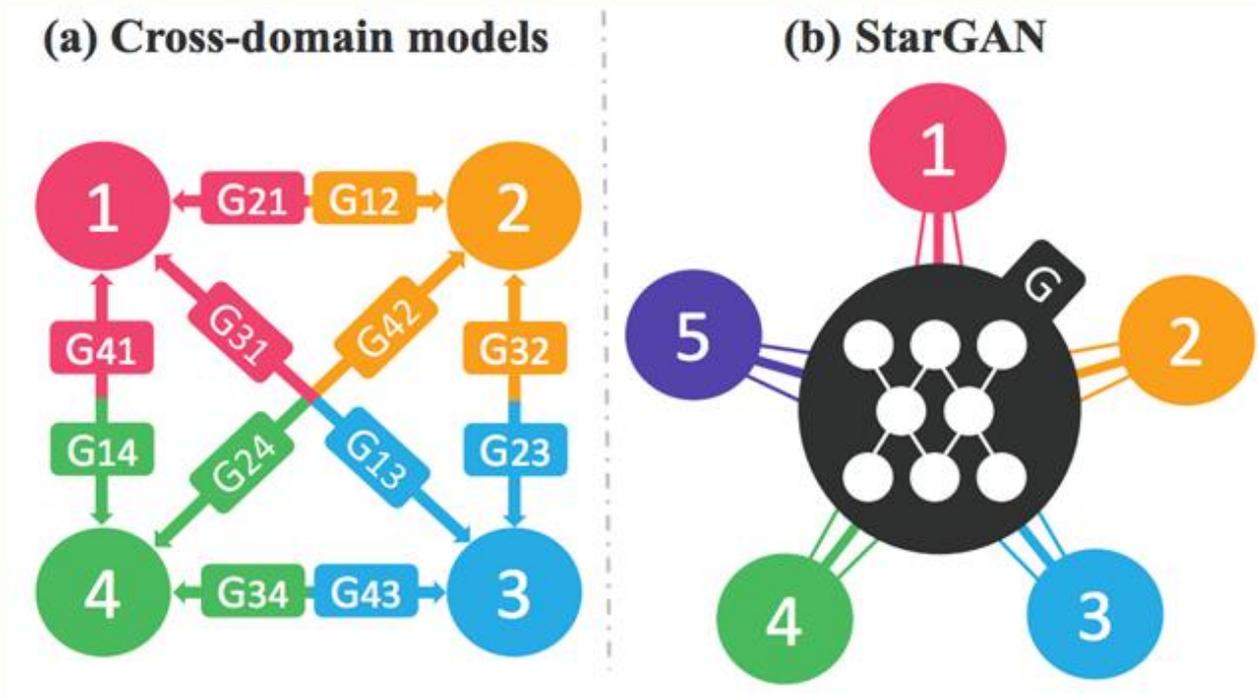
- Conversion is limited to 1 to 1.

→Extend 1 to n



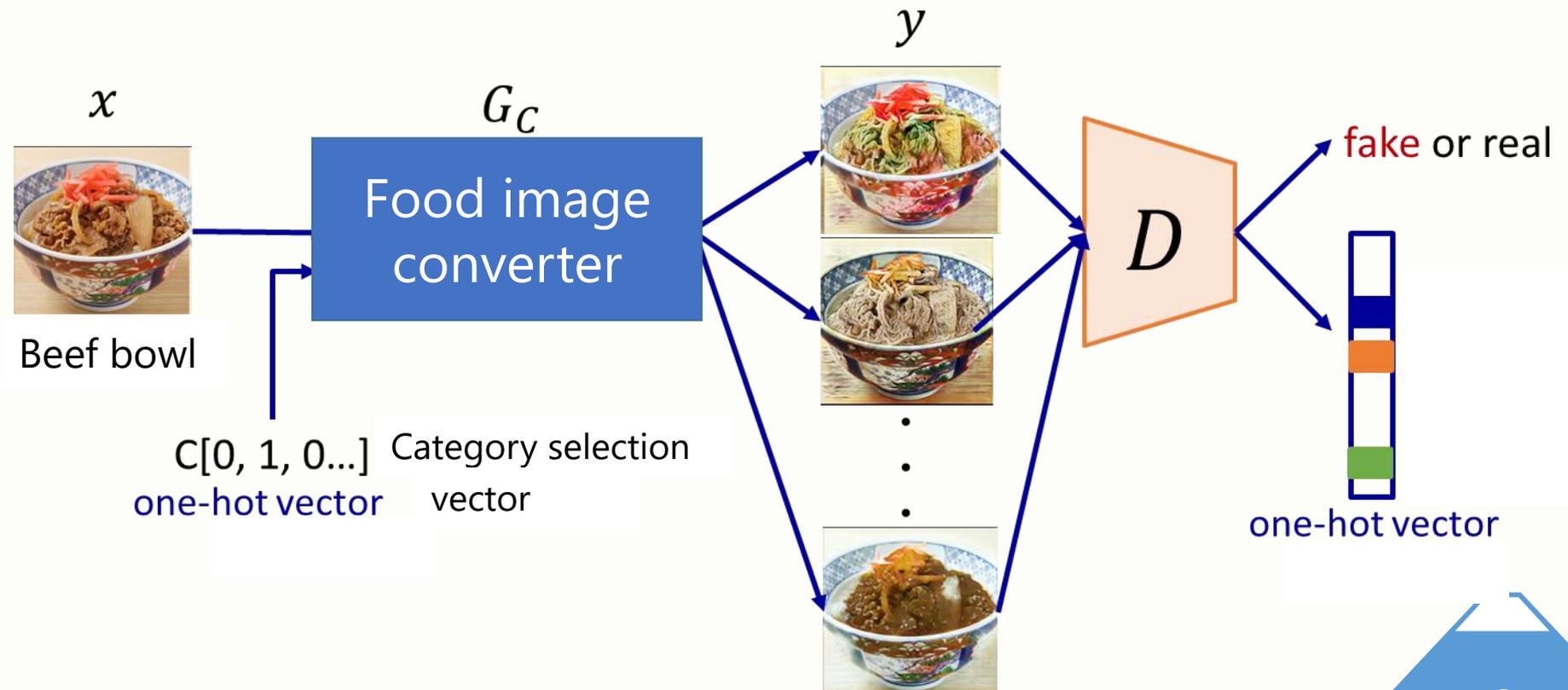
# Conditional CycleGAN

- StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation [Choi+ arXiv-17]
  - Cycle GAN + AC-GAN



# Discriminator for Conditional Cycle GAN

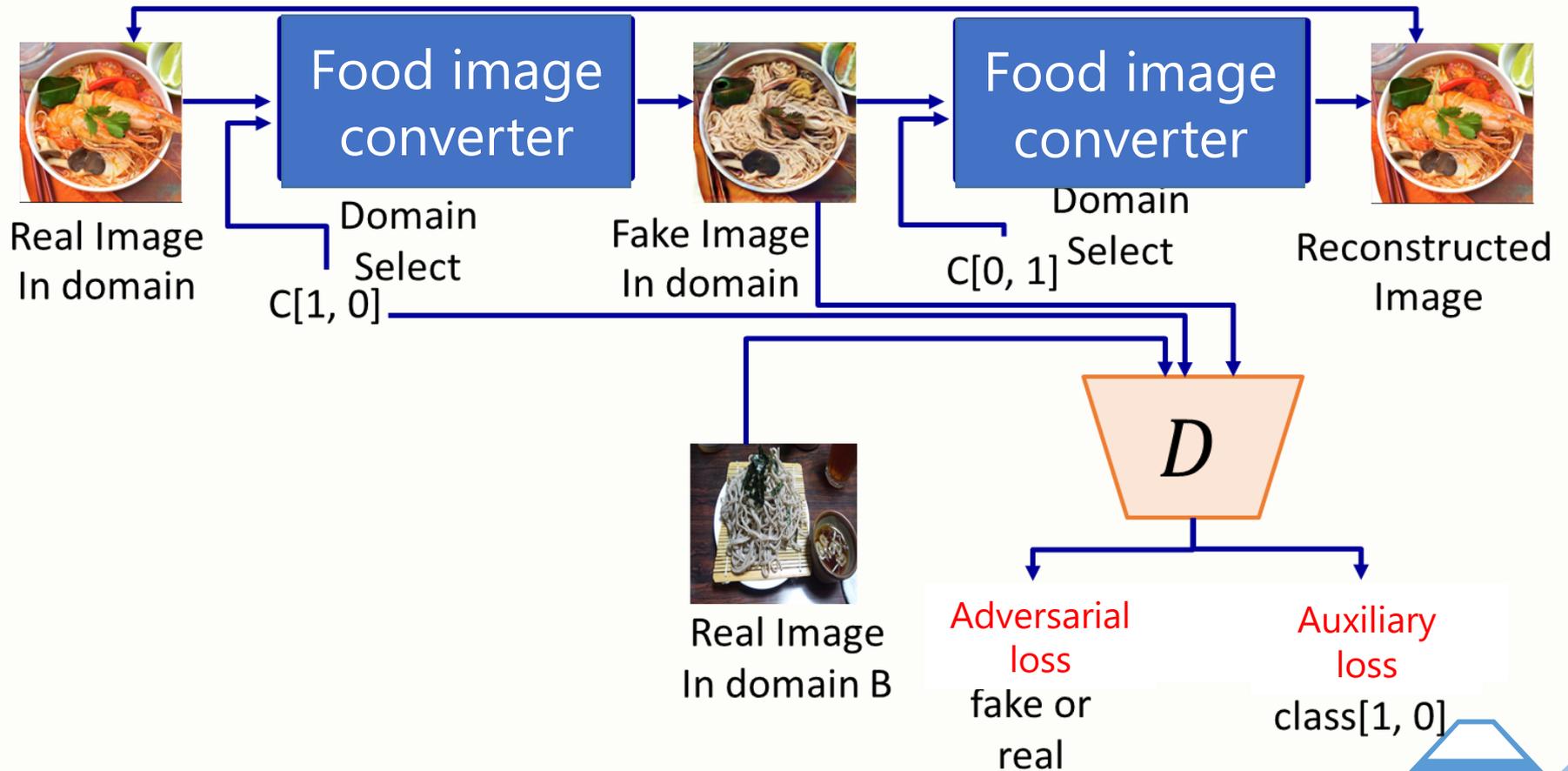
- AC-GAN
- Conditional Image Synthesis With Auxiliary Classifier GANs AC-GAN[Odena+ ICML-17]



# Network Overview

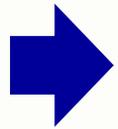
- Network optimizes three types of loss

Consistency Loss



# Experiments

- We use foods, which have similar dish plates as target food category for simplification.



**Selected 10 kinds of food category.**



Curry



Fried rice



Beef bowl



Chilled noodle



Meat spaghetti



Ramen



Rice



Buckwheat noodle



Eel bowl



Fried noodle

# Experimanel Data

- Total amount:
  - 230k images
  - Traing : 0.9
  - Testing : 0.1

Target category	Image number
Chilled noodle	13,499
Meat spaghetti	7,138
Buckwheat noodle	3,530
Ramen	74,007
Fried noodle	24,760
Rice	21,324
Curry rice	34,216
Beef bowl	18,396
Eel bowl	5,329
Fried rice	27,854
total	230,053

# Experimental results

- In case of one food included in an image

**Input**

**Curry**

**Fried rice**

**Beef bowl**

**Chilled  
noodle**

**Meat spaghetti**



# Experimental results

- In case of one food included in an image

**Input**

**Ramen**

**Rice**

**Buckwheat  
noodle**

**Eel bowl**

**Fried noodle**



# Experimental results

- In case of multiple foods included in an image

Input

Curry

Fried rice

Beef bowl

Chilled noodle

Meat spaghetti



# Experimental results

- In case of multiple foods included in an image

input

ramen

rice

Buckwheat  
noodle

Eel bowl

Fried noodle



# Demo video



# Evaluation using user study

2/100



(R)ea (F)ake (S)kip

Done

# Evaluation

	<b>Image generation</b>	<b>Image conversion</b>
Error rate for generated images	12%	30%
Error rate for generated images	16%	32%
Mean error rate	14%	31%

# Conclusion and future work

- Conclusion
  - We transform a food image to another category of a food image automatically
  - We adapted conditional CycleGAN which is an extended version of CycleGAN
- Future work
  - Extend target food categories for conversion.
  - Arbitrary category.