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

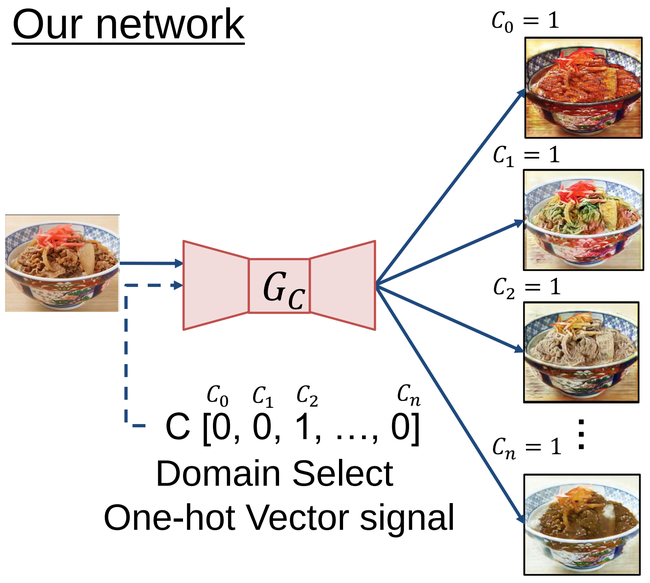
Daichi Horita Ryosuke Tanno Wataru Shimoda Keiji Yanai
The University of Electro-Communications, Tokyo, Japan

Method Overview

Multi-Domain Image-to-Image Translation only to food area

-Based on a Conditional Cycle GAN (cCycle GAN) with a large-scale food image data collected from the Twitter Stream.

Our network



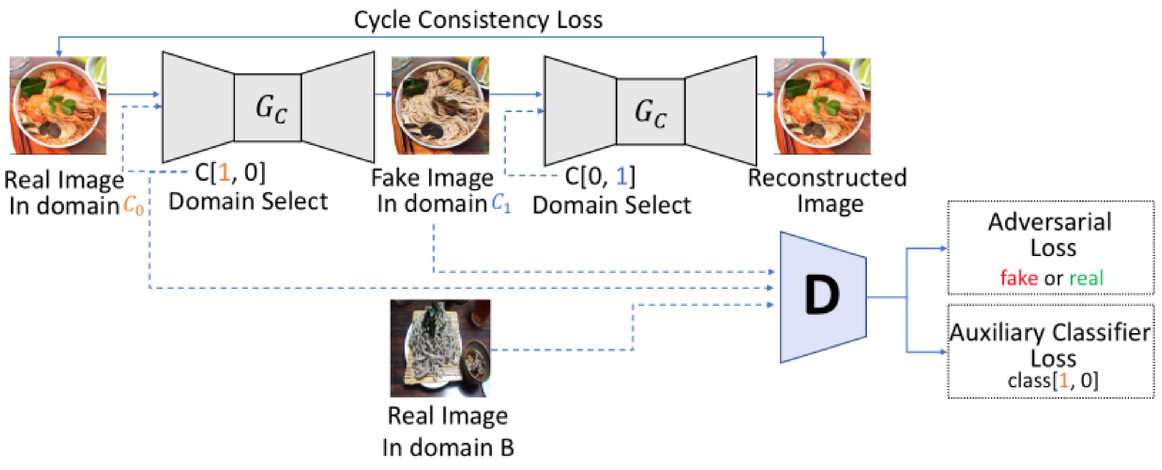
Datasets

-We use Two hundred and thirty thousand food images of ten kinds of typical Japanese foods, but we set restrictions on rice bowl.

category	image number
chilled noodles	13,499
meat spaghetti	7,138
buckwheat noodle	3,530
ramen	74,007
fried noodles	24,760
white rice	21,324
curry rice	34,216
beef bowl	18,396
eel bowl	5,329
fried rice	27,854
TOTAL	230,053

Conditional CycleGAN

Our architecture



-We used StarGAN[1] architecture. Combining loss functions of [2] and [3].

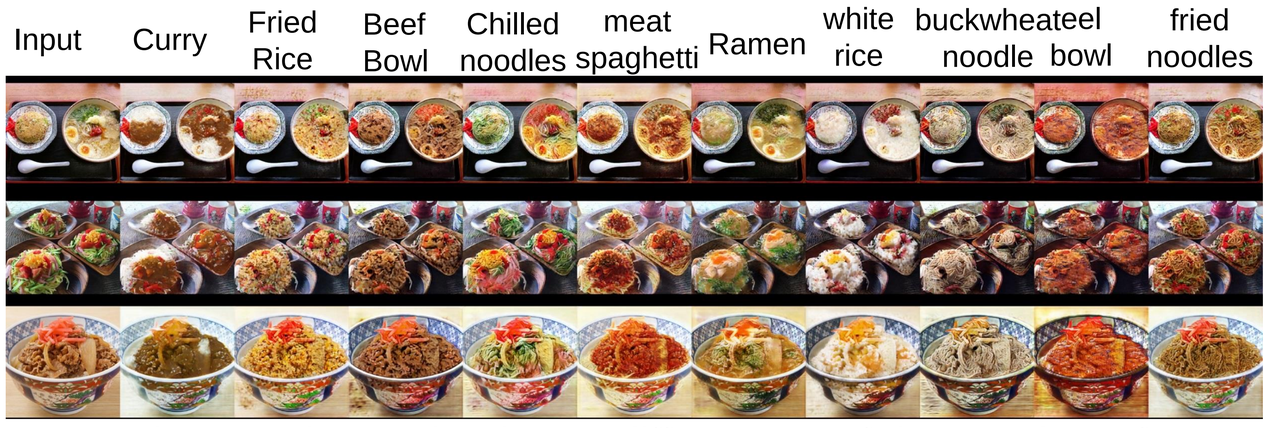
$$L_{acl}^{real} = \mathrm{E}[-logD_{acl}(c'|x)] \qquad L_{acl}^{fake} = \mathrm{E}_{x,c}[-logD_{acl}(c|G(x,c))]$$

$$L_D = L_{adv} + \lambda_{acl}L_{acl}^{real} \qquad L_G = L_{adv} + \lambda_{acl}L_{acl}^{real} + \lambda_{ccl}L_{ccl}$$

$$L_{ccl} = \mathrm{E}_{x,c,c'}||x - G(G(x,c),c')||_1$$

Experiments

☺ We can translate only to food area.



Changes in quality due to differences of the number of total images.

☺ The number of images and the quality are proportional

Future works

-Translation with datasets which excluded constraint of food on rice bowl.

-Translation to categories not included in the datasets.

-Generate higher resolution images.

References

[1] C.Yunjey et al. StarGAN:Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation. CVPR, 2018.

[2] J. Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.

[3] A. Odena et al. Conditional Image Synthesis With Auxiliary Classifier GANs. ICML, 2017

Food Image Generation using A Large Amount of Food Images with Conditional GAN: RamenGAN and RecipeGAN

Yoshifumi Ito Wataru Shimoda Keiji Yanai
The University of Electro-Communications, Tokyo, Japan

Method Overview

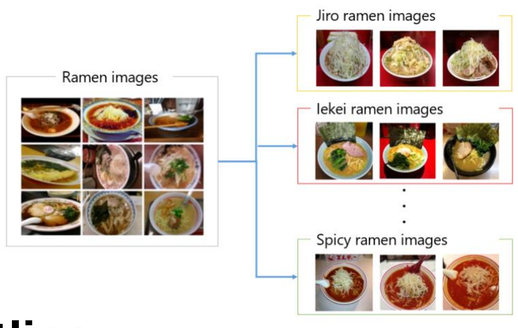
(1) Collecting labeled image.
(2) Conditional Generative Adversarial Network (cGAN).
(3) Improvement of cGAN with a dish plate discriminator and WGAN-GP.
(4) Training and image generation.

Improved Conditional GAN

Coniditonal GAN

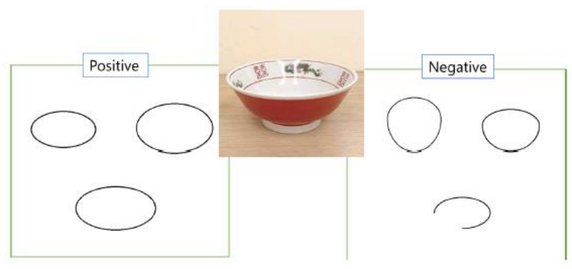
-We used WGAN GP[2] as method of GAN

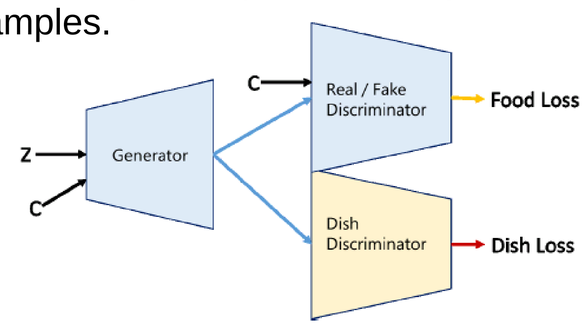
$$\min_G \max_D V(D,G) = \mathrm{E}_{x\sim p_{data}(x)}[logD(x|c)] + \mathrm{E}_{x\sim p_z(x)}[\log(1 - D(x|c))]$$



Discriminator of dish plate outline.

-We prepared an additional discriminator which is trained with oval figure
-Single class classifier with complete oval images as positive samples and the other figure images as negative samples.





Experiments

Ramen GAN

-We trained Conditional GAN (cGAN) model with/without an additional discriminator of dish plates.

Conditional GAN

Improved Conditional GAN

The 6 kind ramen category.

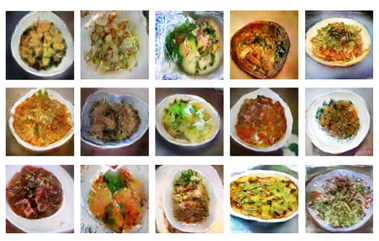
Category	Num
Plane ramen	790
Jiro ramen	5,901
Iekei ramen	2,836
Spicy ramen	3,578
Taiwan ramen	1,567
Onomichi ramen	1,228
TOTAL	15,900

Recipe GAN

-For Recipe GAN, we generated dish images from cooking ingredients.
-Differ from the specific ramen image generation, image generation using recipe data is difficult due to the diversity of the dataset.

Conditional GAN

+WGAN-GP



++Disciriminator of dish plate

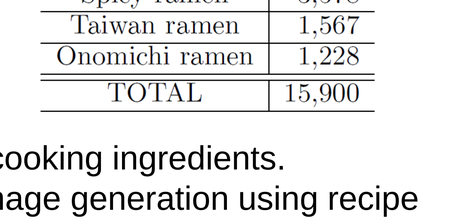
The 10 kind ingredient category.

Category	Num
Onion	29,610
Carrot	22,450
Tomato	18,229
green pepper	8,143
mushroom	7,568

Category	Num
bacon	7,978
red pepper	5986
tofu	9,540
chicken	7,759
pork	10,427
TOTAL	127,690

Recipes retrieval results with generated image

-We show image-based recipe search results which are retrieved by the generated images.



Conditional GAN

Improved Conditional GAN

References

[1] F. Ahmed et al. Improved Training of Wasserstein GANs. NIPS, 2017.

[2] J. Zhu et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV, 2017.