

Magical Rice Bowl: A Real-Time Food Category Changer

Ryosuke Tanno[†] Daichi Horita^{††} Wataru Shimoda^{††} Keiji Yanai^{††}

[†] NTT Communications, ^{††} The University of Electro-Communications, Tokyo, Japan



Project HP

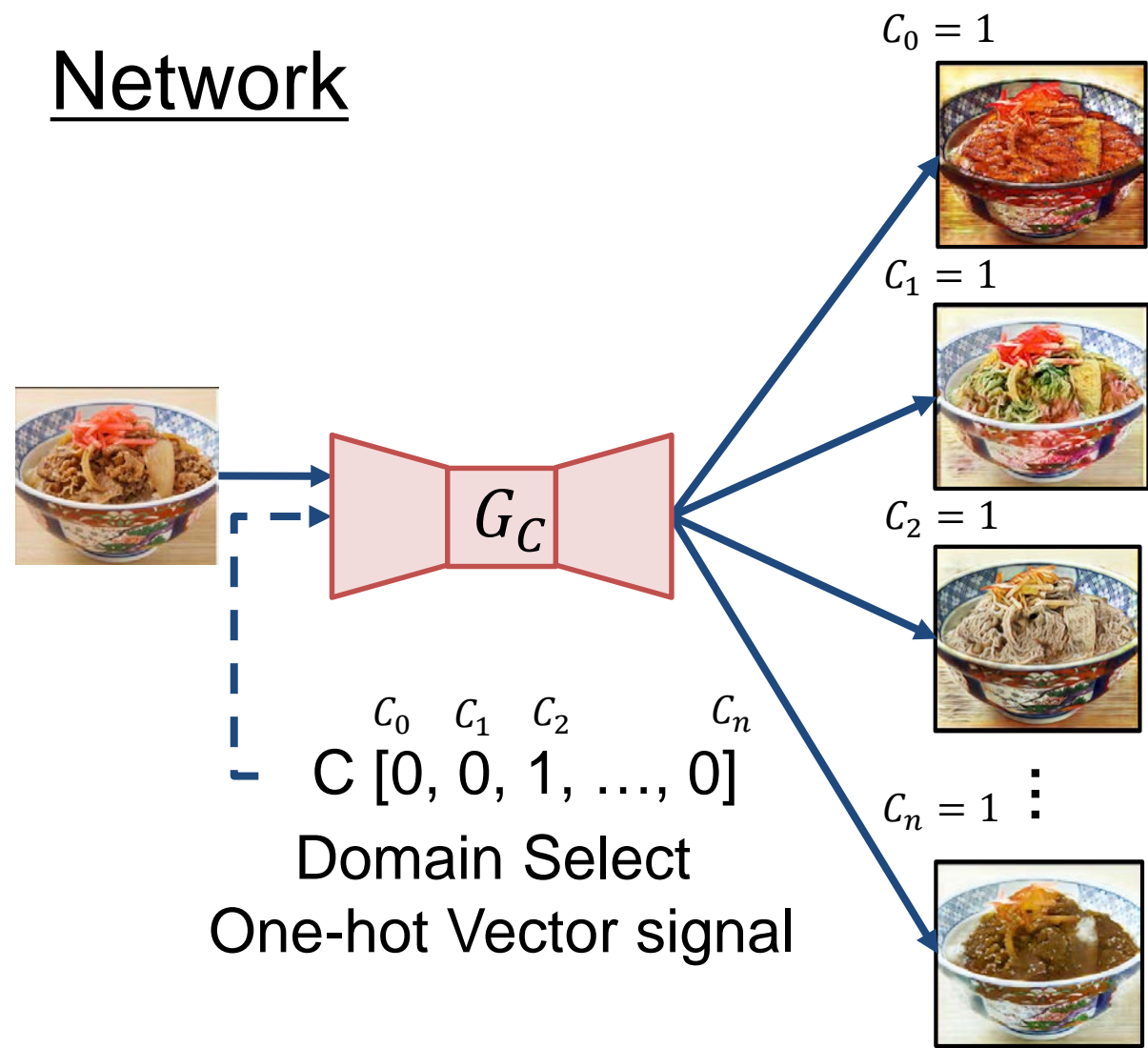


Overview

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

-Based on a StarGAN[1] with a large-scale food image data collected from the Twitter Stream.

Network



-We have gathered 230k food images which consist of 10 kind of food categories from Twitter stream.

Table 1: training data

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

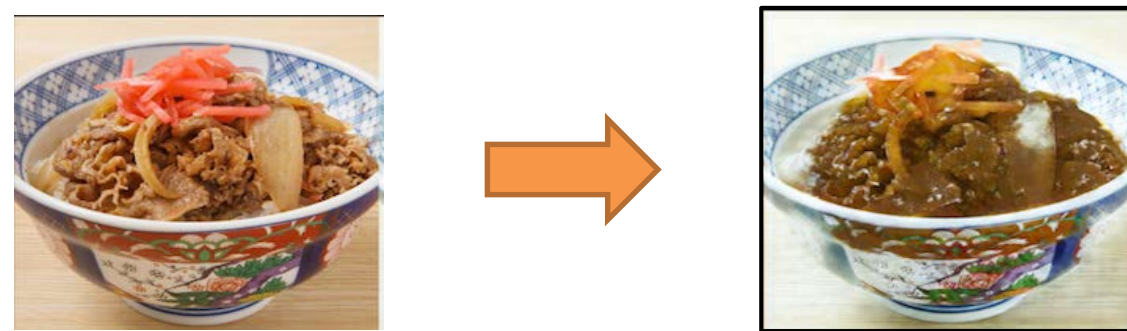
Background

There are many cases in **letters** and **faces** are well generated and translated by GANs.

➔ However, there are few cases of **food**. So we challenged food transformation.

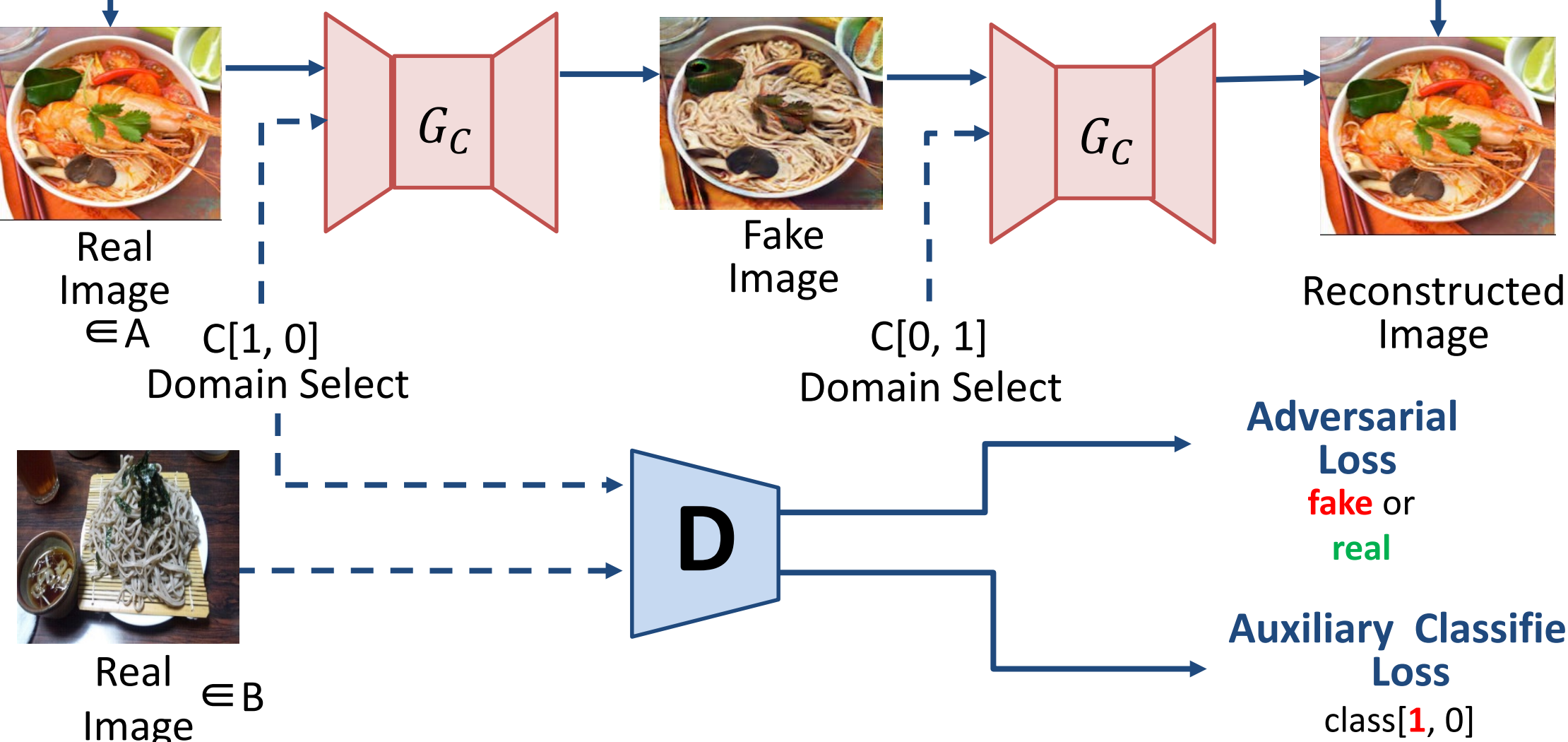


2014 2015 2016 2017
 第 第 第 第 第 第 第 第
 壹 壹 壹 壹 壹 壹 壹 壹
 話 話 話 話 話 話 話 話
 使 使 使 使 使 使 使 使
 徒 徒 徒 徒 徒 徒 徒 徒
 襲 襲 襲 襲 襲 襲 襲 襲
 来 来 来 来 来 来 来 来



Methods

Cycle Consistency Loss



$$L_D = L_{adversarial} + \lambda_{classifier} L_{classifier}$$

$$L_G = L_{adversarial} + \lambda_{classifier} L_{classifier} + \lambda_{cycle} L_{cycle}$$

References

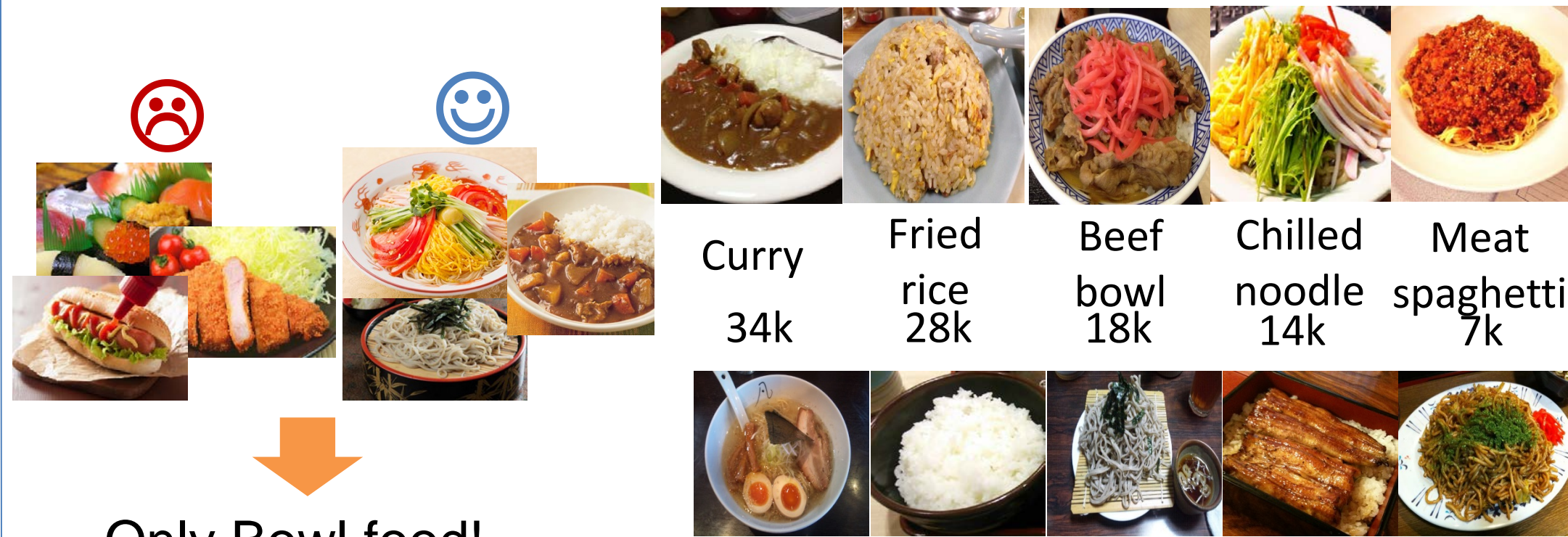
- [1] Y.Choi, M.Choi, M.Kim, J.W.Ha, S.Kim and J.Choo, StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation. In Proc. of CVPR2018.
- [2] J. Y. Zhu, T.Park, P. Isola, A.A. Efros, Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In Proc. of ICCV2017.
- [3] A. Odena, C. Olah, and J. Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. In Proc. of ICML2017.

Experiments

Datasets

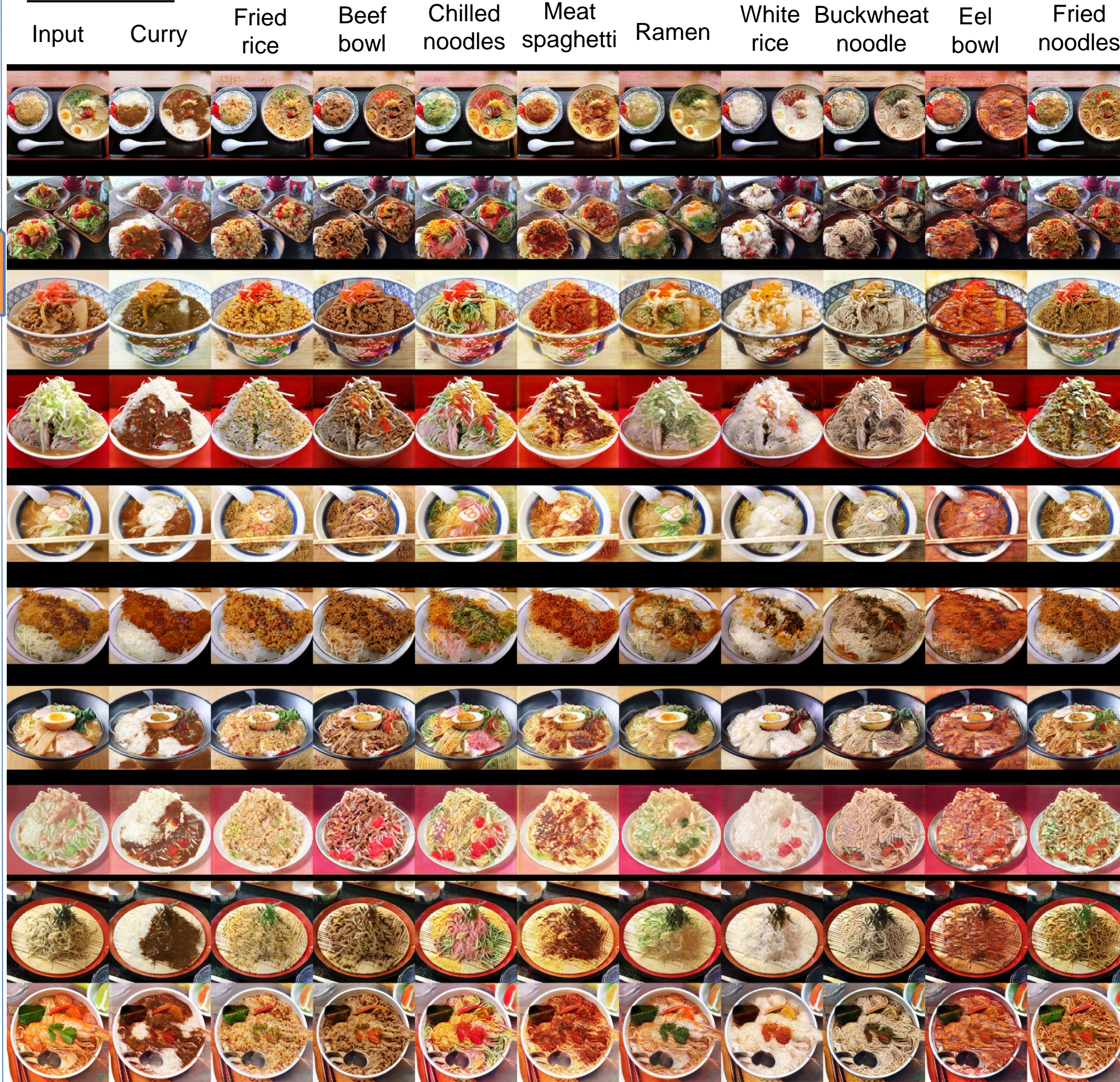
In the case of the dataset has diversity geometric structure, it is difficult to translate.

So we set the constraint of bowl only.



Only Bowl food!

Results



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

😊 The number of images and the quality are proportional



Demo Food

