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

Task

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

Domain transfer (CycleGAN)
CycleGAN

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

Paired
\[ x_i, y_i \]

Unpaired
\[ X, Y \]
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

Food image converter

Corresponding feature “Food”

- Beef bowl to Chilled noodle
- Beef bowl to Buckwheat noodle
- Beef bowl to Curry

...
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
- Adversarial loss
- Auxiliary 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
## Experimantal Data

- **Total amount:**
  - 230k images
  - Training: 0.9
  - Testing: 0.1

<table>
<thead>
<tr>
<th>Target category</th>
<th>Image number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chilled noodle</td>
<td>13,499</td>
</tr>
<tr>
<td>Meat spaghetti</td>
<td>7,138</td>
</tr>
<tr>
<td>Buckwheat noodle</td>
<td>3,530</td>
</tr>
<tr>
<td>Ramen</td>
<td>74,007</td>
</tr>
<tr>
<td>Fried noodle</td>
<td>24,760</td>
</tr>
<tr>
<td>Rice</td>
<td>21,324</td>
</tr>
<tr>
<td>Curry rice</td>
<td>34,216</td>
</tr>
<tr>
<td>Beef bowl</td>
<td>18,396</td>
</tr>
<tr>
<td>Eel bowl</td>
<td>5,329</td>
</tr>
<tr>
<td>Fried rice</td>
<td>27,854</td>
</tr>
<tr>
<td><strong>total</strong></td>
<td><strong>230,053</strong></td>
</tr>
</tbody>
</table>
Experimental results

- In case of one food included in an image
Experimental results

- In case of one food included in an image

<table>
<thead>
<tr>
<th>Input</th>
<th>Ramen</th>
<th>Rice</th>
<th>Buckwheat noodle</th>
<th>Eel bowl</th>
<th>Fried noodle</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Image of Ramen" /></td>
<td><img src="image2" alt="Image of Ramen" /></td>
<td><img src="image3" alt="Image of Rice" /></td>
<td><img src="image4" alt="Image of Buckwheat noodle" /></td>
<td><img src="image5" alt="Image of Eel bowl" /></td>
<td><img src="image6" alt="Image of Fried noodle" /></td>
</tr>
</tbody>
</table>
Experimental results

- In case of multiple foods included in an image
Experimental results

- In case of multiple foods included in an image

<table>
<thead>
<tr>
<th>input</th>
<th>ramen</th>
<th>rice</th>
<th>Buckwheat noodle</th>
<th>Eel bowl</th>
<th>Fried noodle</th>
</tr>
</thead>
</table>

- 2018/7/13
- UEC yanailab 2017
- 15
Demo video
Evaluation using user study

2/100

(R)eal  (F)ake  (S)kip

Done
## Evaluation

<table>
<thead>
<tr>
<th></th>
<th>Image generation</th>
<th>Image conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error rate for generated images</td>
<td>12%</td>
<td>30%</td>
</tr>
<tr>
<td>Error rate for generated images</td>
<td>16%</td>
<td>32%</td>
</tr>
<tr>
<td>Mean error rate</td>
<td>14%</td>
<td>31%</td>
</tr>
</tbody>
</table>

2018/7/14  UEC yanailab 2017
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.