Ramen Spoon Eraser: CNN-based photo transformation for improving attractiveness of ramen photos

Daichi Horita, Jaehyeong Cho, Takumi Ege, Keiji Yanai (The University of Electro-Communications, Tokyo)

Introduction

- A ramen spoon in a bowl spoils its appearance greatly as shown in the right figure.
- To improve attractiveness of ramen photos, erasing ramen spoon is indispensable.
- High-quality image transformation is possible with GAN-based image-to-image transformation (Pix2pix) [3], which potentially erases ramen spoons from photos.

Proposed Method

- We adopt a U-Net [2] as a image transformation network.
- We train it in two ways: MSE loss and Pix2pix loss [3]
- Mean Squared Error (MSE) Loss

\[ \ell(x, y) = L = \{l_1, \ldots, l_N\}^T, \quad l_n = (x_n - y_n)^2, \]

- Pix2pix [3] (with adversarial training)

\[ \mathcal{L}_{G_{\text{GAN}}}(G, D) = \mathbb{E}_{x, y} \left[ \log D(x, y) \right] + \mathbb{E}_{x, y} \left[ \log (1 - D(x, G(x, z))) \right]. \]

\[ \mathcal{L}_{L_1}(G) = \mathbb{E}_{x, y, z} \left[ \|y - G(x, z)\|_1 \right]. \]

\[ G^* = \min_G \max_D \mathcal{L}_{G_{\text{GAN}}}(G, D) + \lambda \mathcal{L}_{L_1}(G). \]

Experimental Results

- [Training data] 6000 pairs of ramen-without/with-spoon photos
  - We generated ramen-with-spoon photos by overlaying ramen spoon images on the ramen-without-spoon photos randomly.
  - We trained two model using the MSE loss and the Pix2pix loss.
  - We compare two models:
    - MSE loss erased spoons. However, the results are blurred.
    - Pix2pix erased spoons more naturally than MSE loss.
  - In the third row, both models failed to remove spoons, because the training data did not contain enough images with red ramen spoons.
  - We have not evaluated which one seems to be more attractive by subjects.
    - So, we would like you to attach a sticker to more attractive images!!

<table>
<thead>
<tr>
<th>MSE Loss</th>
<th>Pix2pix</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Conclusions

- We proposed an application which erases a spoon from a ramen-with-spoon photo by CNN-based image transformation.
- We plan to extend our model for diverse utensils such as chopsticks, forks and knives to make food photos more attractive.

References