Ketchup GAN: A New Dataset for Realistic Synthesis of Letters on Food

Gibran Benitez-Garcia, and Keiji Yanai

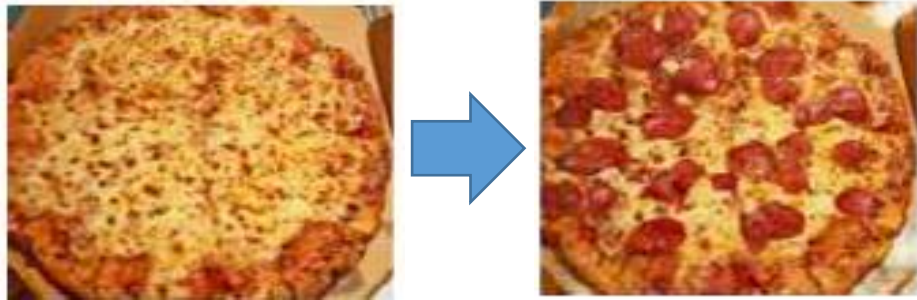The University of Electro-Communications
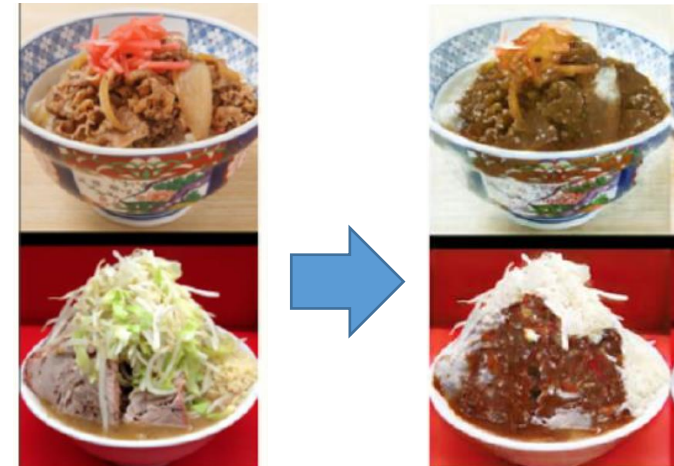
MMArt-ACM '21   2021.11.16

Motivation

– **Food image generation and translation** have been attracted particular attention due to the outstanding performance of GAN networks

– To train GAN models, a **large amount of real-world images** is needed

– There is no GAN-based model proposed for **food image decoration**

PizzaGAN[*]

FoodGAN[**]

* "How to make a pizza: Learning a compositional layer-based GAN model," CVPR **2019**

** "Unseen Food Creation by Mixing Existing Food Images with Conditional StyleGAN," MADiMa **2019**

Objective

- **_Objective:_**
  - _Introduce a new dataset for realistic synthesis of letters on food_

- **_Approach:_**
  - _Gathering real-world images of omelettes decorated with ketchup letters and generate egg and ketchup masks, and synthetic ketchup-free images_

Original photos

Egg & Ketchup masks

Synthetic ketchup-free

The Ketchup GAN Dataset

– *We gathered **omelette rice images** with and without ketchup letters **from Twitter*** in Japan (part of UECFOOD-100 dataset)*

Images gathered from Twitter

Geotagged Twitter photo mining from UECFOOD-100

* "Large-scale Twitter Food Photo Mining and Its Applications," BigMM *2019*

The Ketchup GAN Dataset

- *We train a CNN classifier with meaningful omelette dishes, **five classes are obtained with 2000** images per class*

- *We train a **weakly supervised segmentation*** algorithm for obtaining ketchup and omelette masks*

Early results from weakly supervised semantic segmentation

* "Learning pixel-level semantic affinity with image-level supervision for weakly supervised semantic segmentation," CVPR *2018*

– *Four sets with 1,150 images* each conform the complete dataset

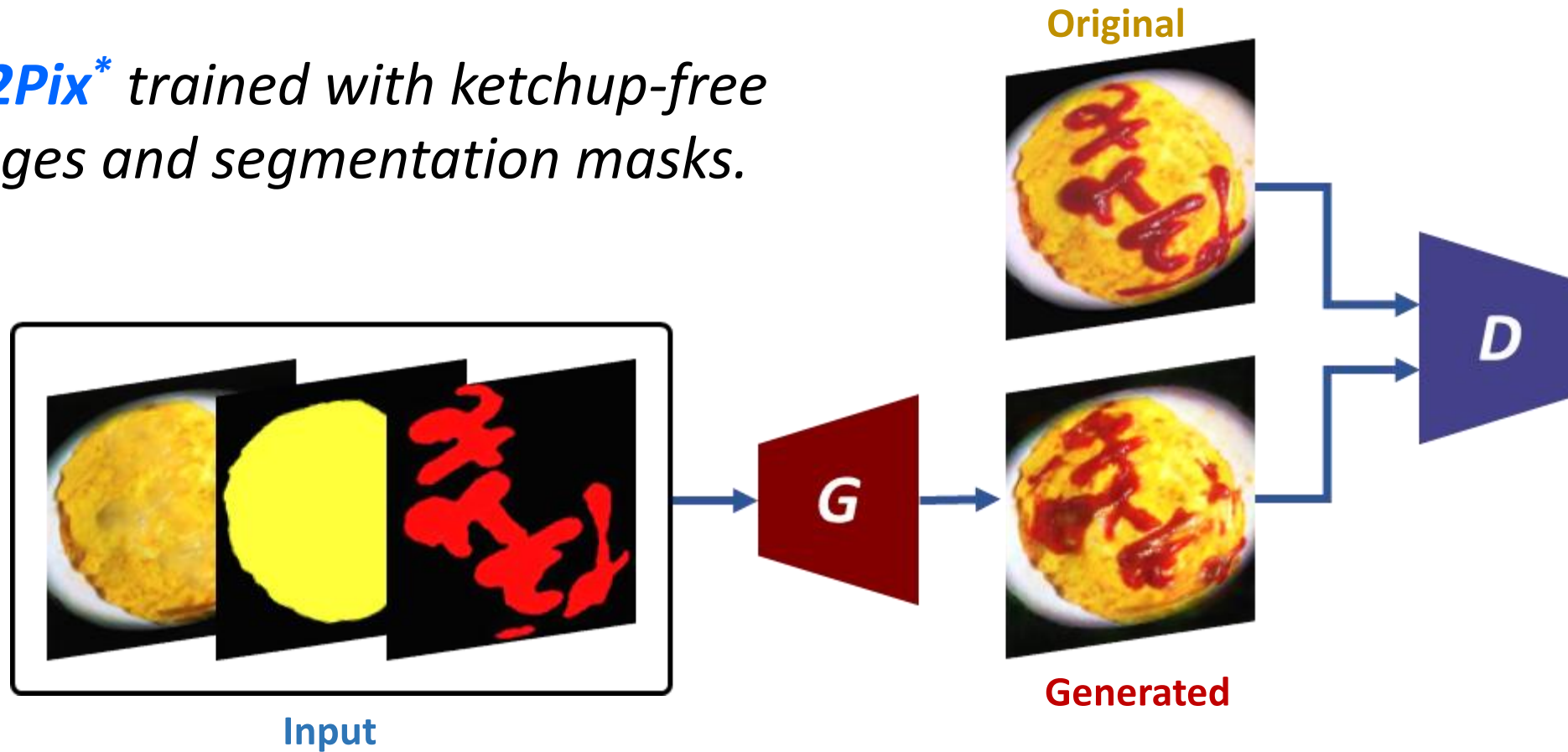**Omelette with ketchup**   **Segmentation masks**   **Synthetic ketchup-free**   **Original plain omelette**
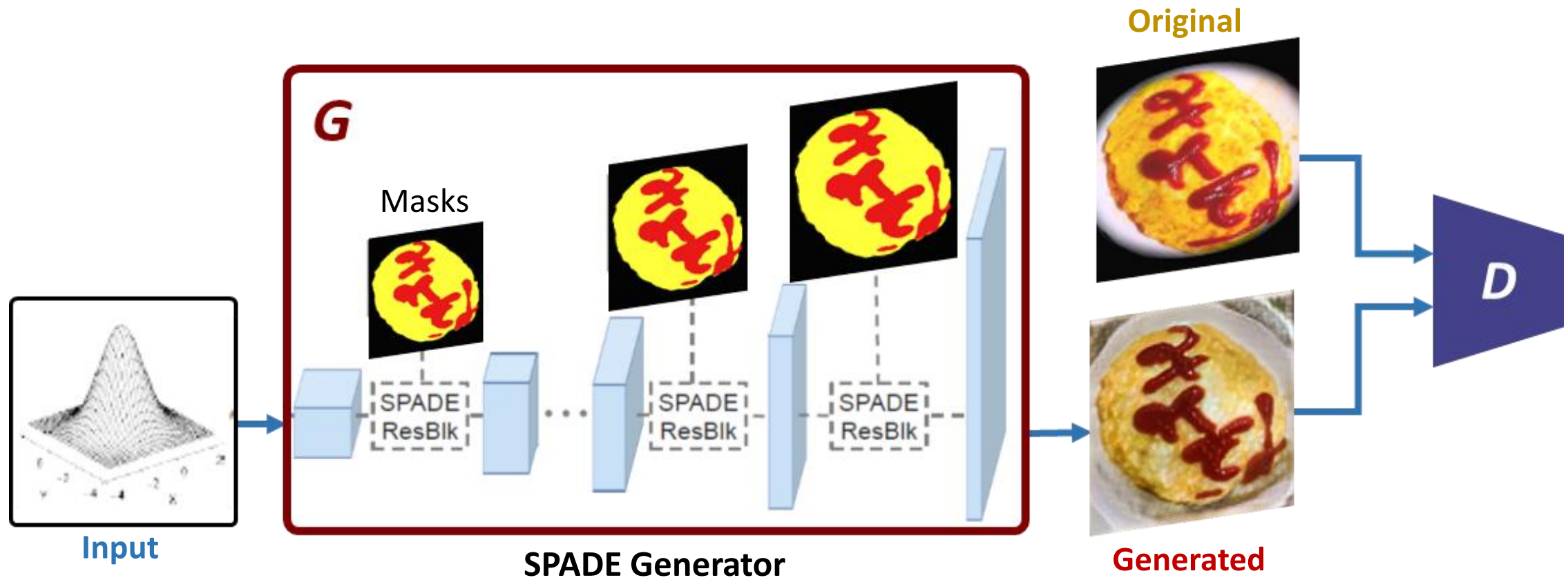
Photorealistic Ketchup Letter Synthesis

– *We **test two GAN models** with our dataset*

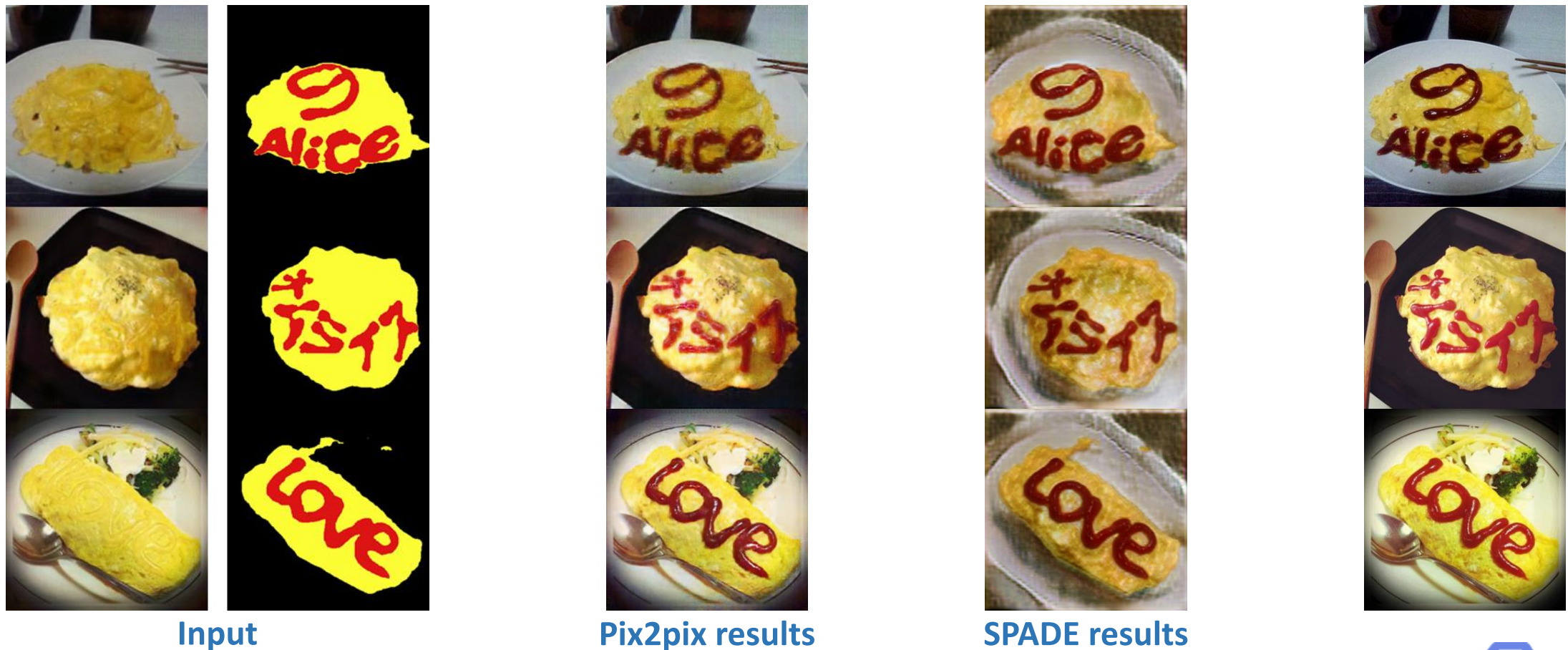– ***Pix2Pix*** trained with ketchup-free images and segmentation masks.*

Original

Generated

Input

* "Image-to-image translation with conditional adversarial networks," CVPR *2017*

Photorealistic Ketchup Letter Synthesis

– **SPADE*** *trained only with segmentation masks.*

* "Semantic image synthesis with spatially-adaptive normalization," CVPR *2019*

Photorealistic Ketchup Letter Synthesis

– *Results from test set* of our dataset.

Original

Input   **Pix2pix results**   **SPADE results**

Photorealistic Ketchup Letter Synthesis

– *Results with **customized ketchup patterns**.*

Pattern   **Input**   **Pix2pix results**   **SPADE results**
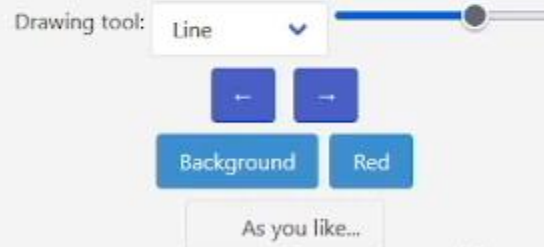
Omelette Decoration (Application)

* "Fast and Accurate Real-Time Semantic Segmentation with Dilated Asymmetric Convolutions," ICPR *2020*
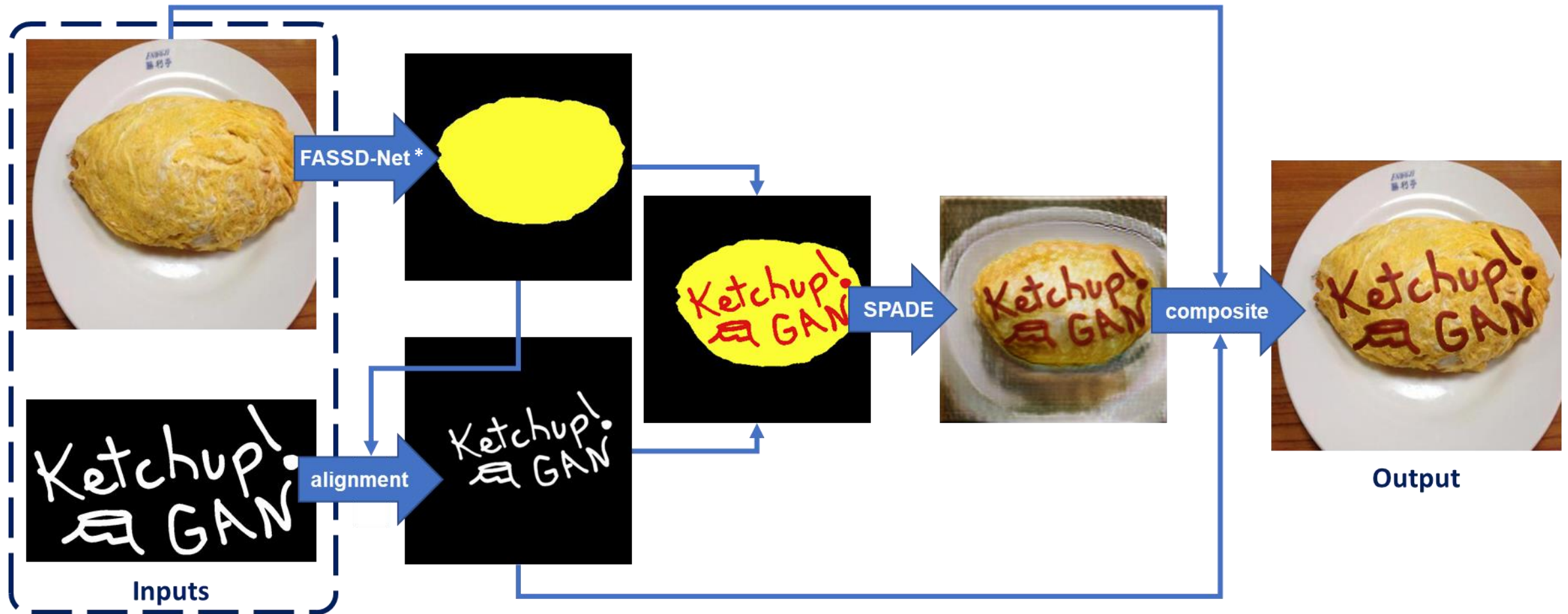
– *Comparison with **manually edited** photo.*

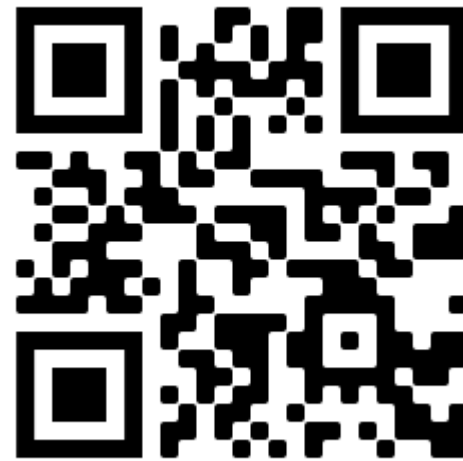# Web Demo

Conclusion

- ***Conclusions:***

  - *We introduced a new dataset for the realistic synthesis of ketchup.*

  - *We evaluated two GAN-based methods for ketchup synthesis using our dataset, and proposed a simple omelette decoration application.*

  - *Our results show that the proposed dataset is suitable for training GAN-based methods.*

- ***Future work:***

  - *Improve the application by proposing a GAN model capable of aligning the input texts and synthesizing omelettes with ketchup patterns simultaneously.*

Thank You For Your Attention

http://mm.cs.uec.ac.jp/omrice/