MIAIS: A Multimedia Recipe Dataset with Ingredient Annotation at Each Instructional Step

Yixin Zhang[1], Yoko Yamakata[2] and Keishi Tajima[1]

[1]Kyoto University and [2]The University of Tokyo, Japan

KYOTO UNIVERSITY

THE UNIVERSITY OF TOKYO
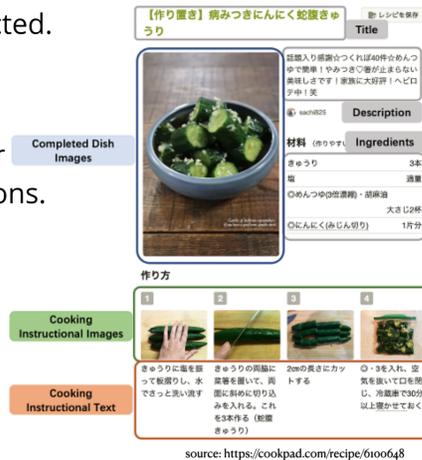
RESEARCH BACKGROUND

More and more datasets on recipes are constructed.

Food-101 dataset, Recipe1M+, UECFood256:
- mainly focus on **completed dish images** after cooking rather than images in recipe instructions.

Cookpad dataset:
- 1, 642, 450 images of completed dishes
- 3, 105, 594 **instructional images**

source: https://cookpad.com/recipe/6100648

RESEARCH PROBLEM AND PURPOSE

Most of the datasets:
- No sufficient **multimedia information (text and image) on the cooking instructions** of the recipe.

However, the **instructional information** is one of the most important and unique characteristics of recipe data.

MIAIS Dataset
- *Goal*: construct a recipe dataset with sufficient **multimedia data** and the **annotations** to them for every cooking step (based on Cookpad dataset)
- *Applications*: cooking flow graph generation, recipe text generation, and cooking action recognition

ANNOTATION AND COOKING FLOW GRAPH

Challenge and Common Solutions

The original form sometimes makes it difficult to understand recipes directly and efficiently.

Recipe form: Sequential

Actual actions: sometimes parallel
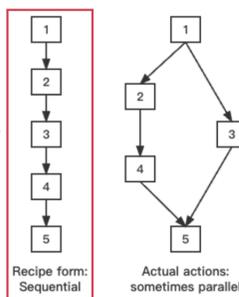
- A <u>challenge</u> in understanding the cooking process from recipes:
  ○ present the complicated operations of cooking activities in a succinct workflow.

- A **common solution** is to **generate a flow graph** to help them understand the cooking process.
  ○ Focus on the flow of ingredients in recipes.

- However, the ingredient information in text tends to be insufficient or omitted when written by users.

Our new annotation rule

Each instructional step is annotated with
all ingredients that physically exist in the intermediate products processed in the step
no matter whether they are mentioned in the instructional text or not.

Flow Graph Generation

- We **track the paths of ingredients** in order to generate the cooking flow graph.

- If a specific ingredient
  ○ is contained in a previous step;
  ○ can also be detected in one of the subsequent steps
- we can consider that there is continuity between the two steps.
- The path is created for each ingredient and the final flow graph is a union set of all the paths.

- E.g.,
- "**Potato**": 1-2-3-4-7-8;
  "**Beef**": 3-5-6-7-8

The Original Multimedia Recipe
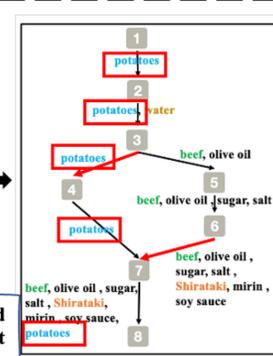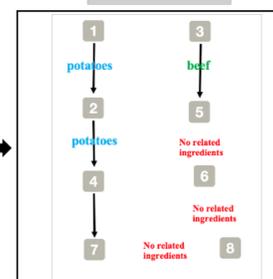
Our proposed annotation method

Our method complement those omitted information

Cooking flow graph

- The flow graph of this recipe can be generated based on the above tracking information.
- The order in which the actual cooking is performed is **not sequential, but with several branches**.
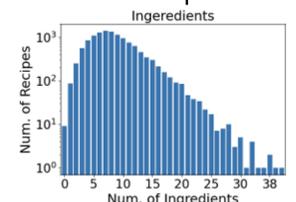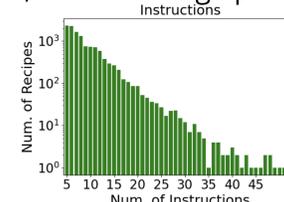
MIAIS DATASET

Multimedia recipe data
12,000 recipes in Japanese
The text and image data are originated from Cookpad.
- Recipe selection **rule**:
  ○ Recipes that have at least **five** cooking steps
  ○ Recipes that contain a text description and an image for **every step**

- 109,547 text-image paired instructional steps in total

CONCLUSION

- A multimedia recipe dataset with ingredient annotation for every instructional step
  ○ Containing **both text and image data for every cooking step**
  ○ Supplementary information about text-image pair of every step

- Cooking flow graphs
  ○ Directly derived based on the annotation information
  ○ These flow graphs can become the ground truth for the method of recipe flow graph generation