Food Scanner Application - Inverse Cooking System in Backend

Anusree Jnanakrishnan Asst.Prof. Hussain Ahmed

MCA, Semester III
Department of Computer Applications
Government Engineering College, Thrissur

20MC245 - Mini Project First Review, 2022

Outline

Introduction

- Deep learning can be defined as the method of machine learning and artificial intelligence that is intended to intimidate humans and their actions based on certain human brain functions to make effective decisions. In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound.
- A convolutional neural network is a specific kind of neural network with multiple layers. It processes data that has a grid-like arrangement then extracts important features. One huge advantage of using CNNs is that you don't need to do a lot of pre-processing on images.
- The goal of this project was to use the largest publicly available collection of recipe data (Recipe1M+) to build a Food Scanning Application for ingredients and recipes. Train, evaluate and test a model able to predict title, ingredients and recipe. Finally, to build Food Scanning application.

Existing System

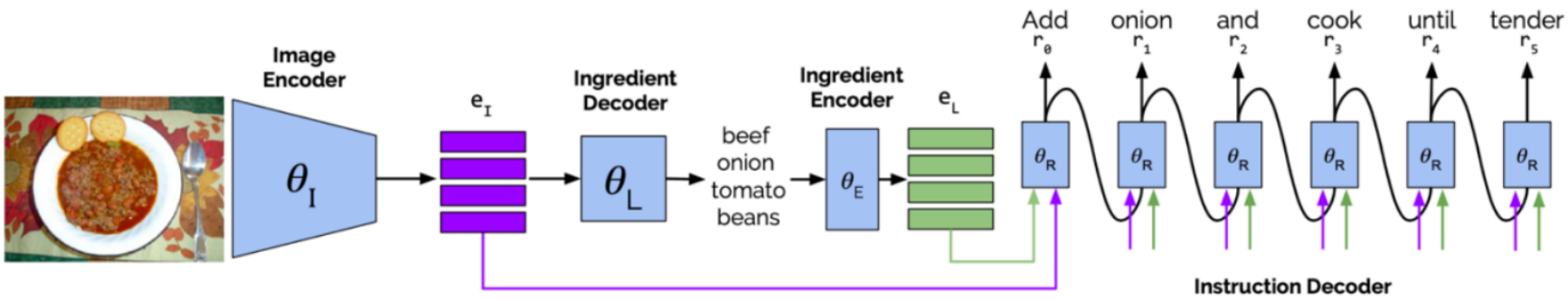
- In the past, algorithms have been using simple systems of recipe retrieval based on image similarities in some embedding space. This approach is highly dependent on the quality of the learned embedding, dataset size and variability.
- Therefore, these approaches fail when there is no match between the input image and the static dataset .
- Inverse cooking algorithm instead of retrieving a recipe directly from an image, proposes a pipeline with an intermediate step where the set of ingredients is first obtained.
- This allows the generation of the instructions not only taking into account the image, but also the ingredients.

Proposed System

- Recipe1M affords the ability to train high-capacity models on aligned, multi-modal data. Using these data, train a neural network to learn a joint embedding of recipes and images that yields impressive results on an image-recipe retrieval task.
- The Food Scanner Application which works with the help of 1M+ Dataset which works in the backend.
- The user can easily access the details of scanned food item.

Architecture

- Inverse Cooking recipe generation model with the multiple encoders and decoders, generating the cooking instructions



Modules

- MODULE 1
  - Encoder
  - 1.1 Image Encoder
  - 1.2 Ingredient Encoder
- MODULE 2
  - Decoder
  - 2.1 Instruction Decoder
  - 2.1 Ingredient Decoder

Module Description

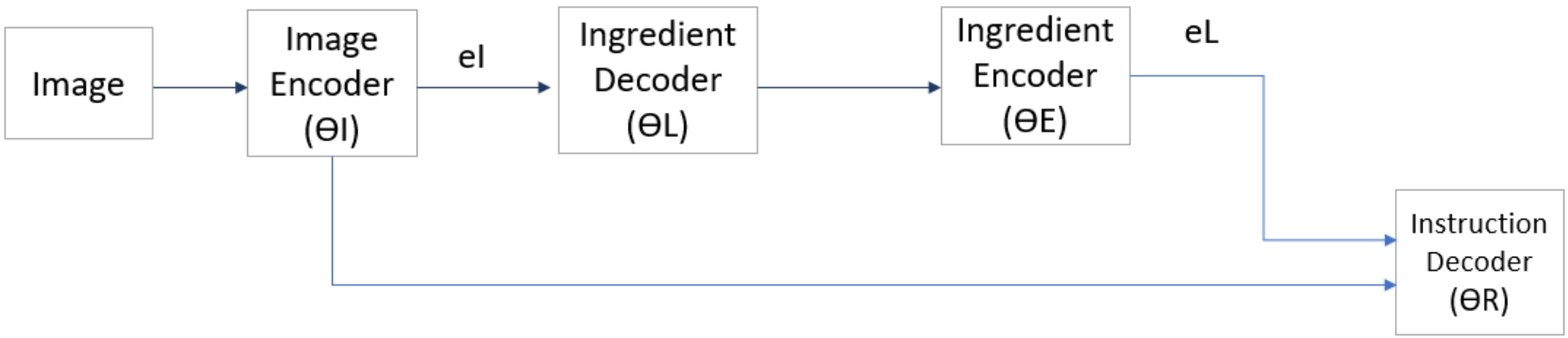
- Image Encoder - We extract image features eI with the image encoder, parametrized by θI .
- Ingredient Decoder - Ingredients are predicted by θL
- Ingredient Encoder - encoded into ingredient embeddings eL with θe.
- Cooking instruction decoder - parametrized by θR generates a recipe title and a sequence of cooking steps by attending to image embeddings eI , ingredient embeddings eL, and previously predicted words (r0, ..., rt1).

Algorithm

- This recipe retrieval algorithm was developed by the Facebook AI Research and it is able to predict ingredients, cooking instructions and a title for a recipe, directly from an image .
- We applied CNN to the tasks of food detection and recognition through parameter optimiza-tion. We constructed a dataset of the most frequent food items in a publicly available 1M+ Dataset, and used it to evaluate recognition performance.
- PyTorch : All models are implemented with PyTorch. The greedy search function implements Greedy Search, which simply picks the most likely token at every step. This is the fastest and simplest algorithm, but can work well if the model is properly trained.
- Resnet50 is a pre-trained Deep learning model. A pre-trained model is trained on a different task than the task at hand and provides a good starting point since the features learned on the old task are useful for the new task.

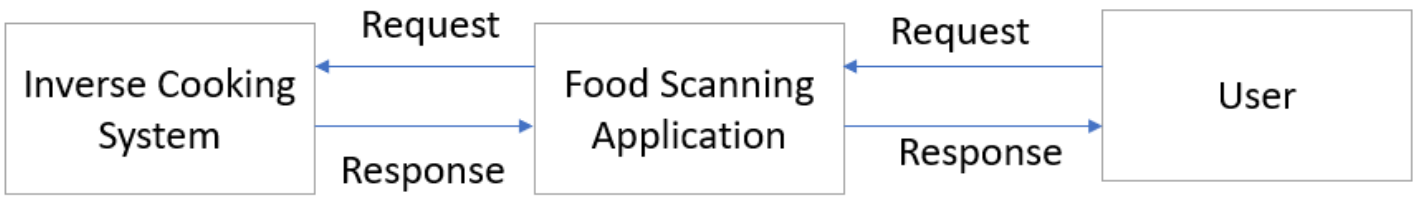
Data Flow Diagram I

- 1.Inverse Cooking System(Backend Working)

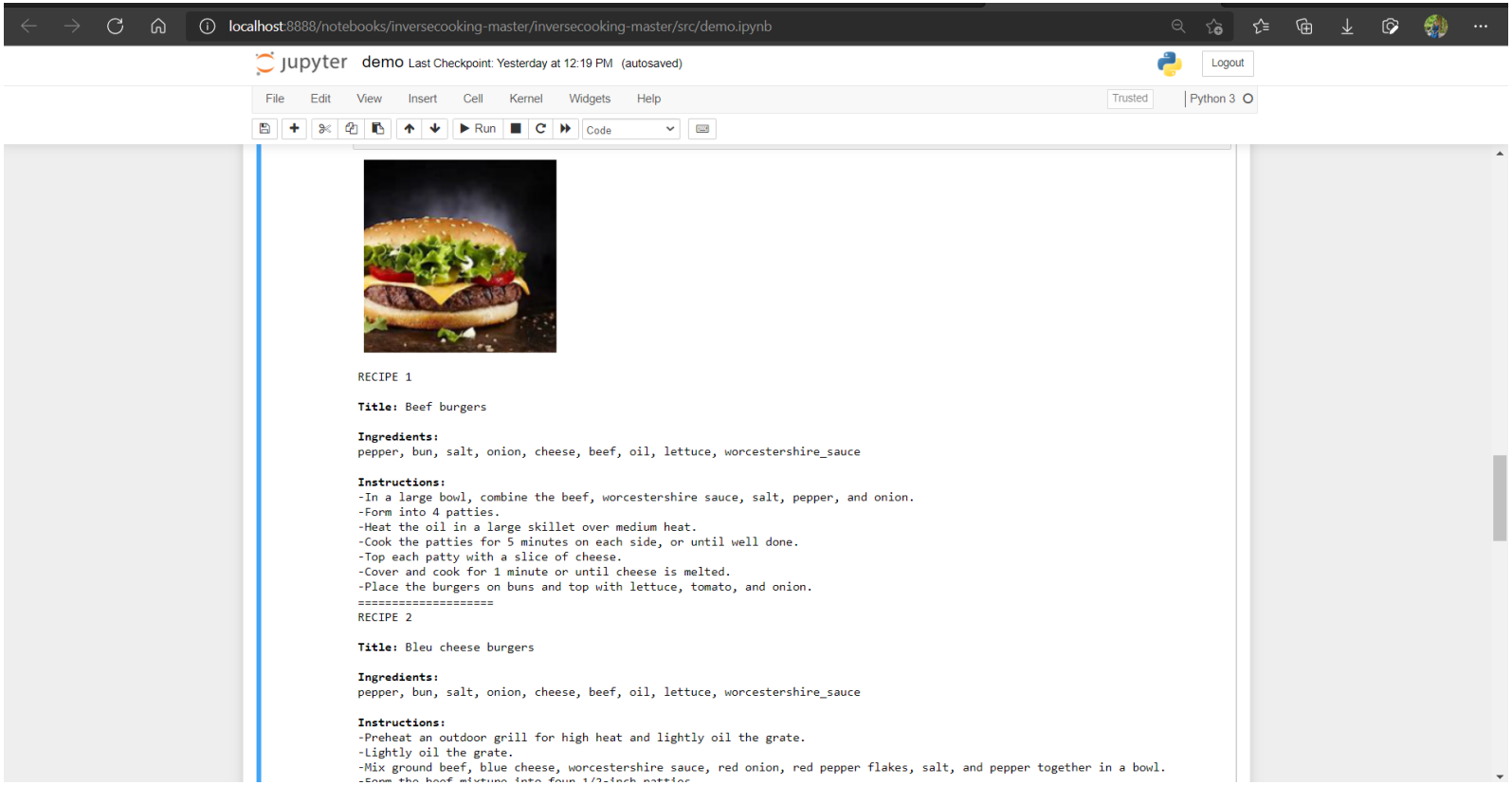


Data Flow Diagram II

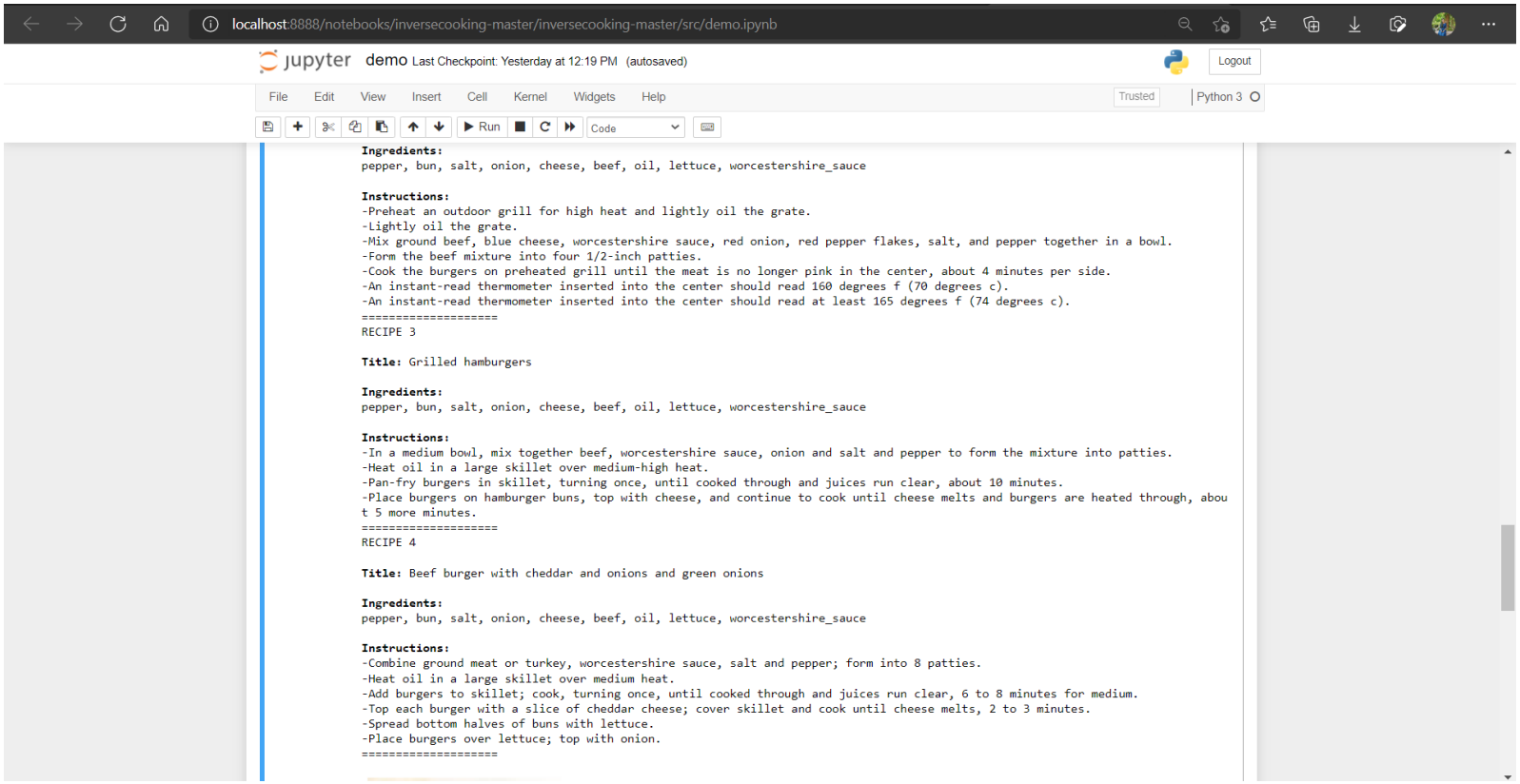
- 2. Food Scanner



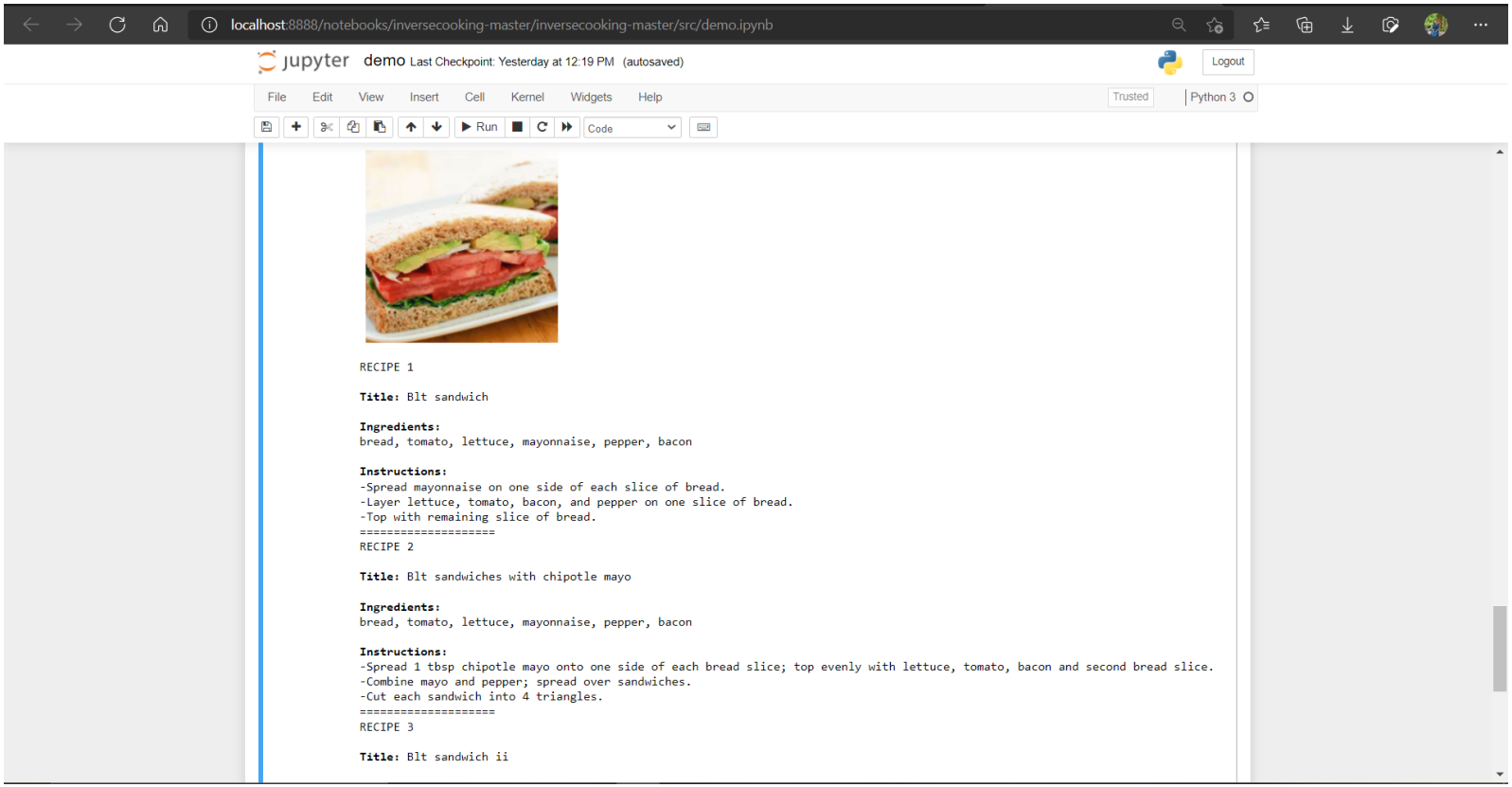
Sample Result



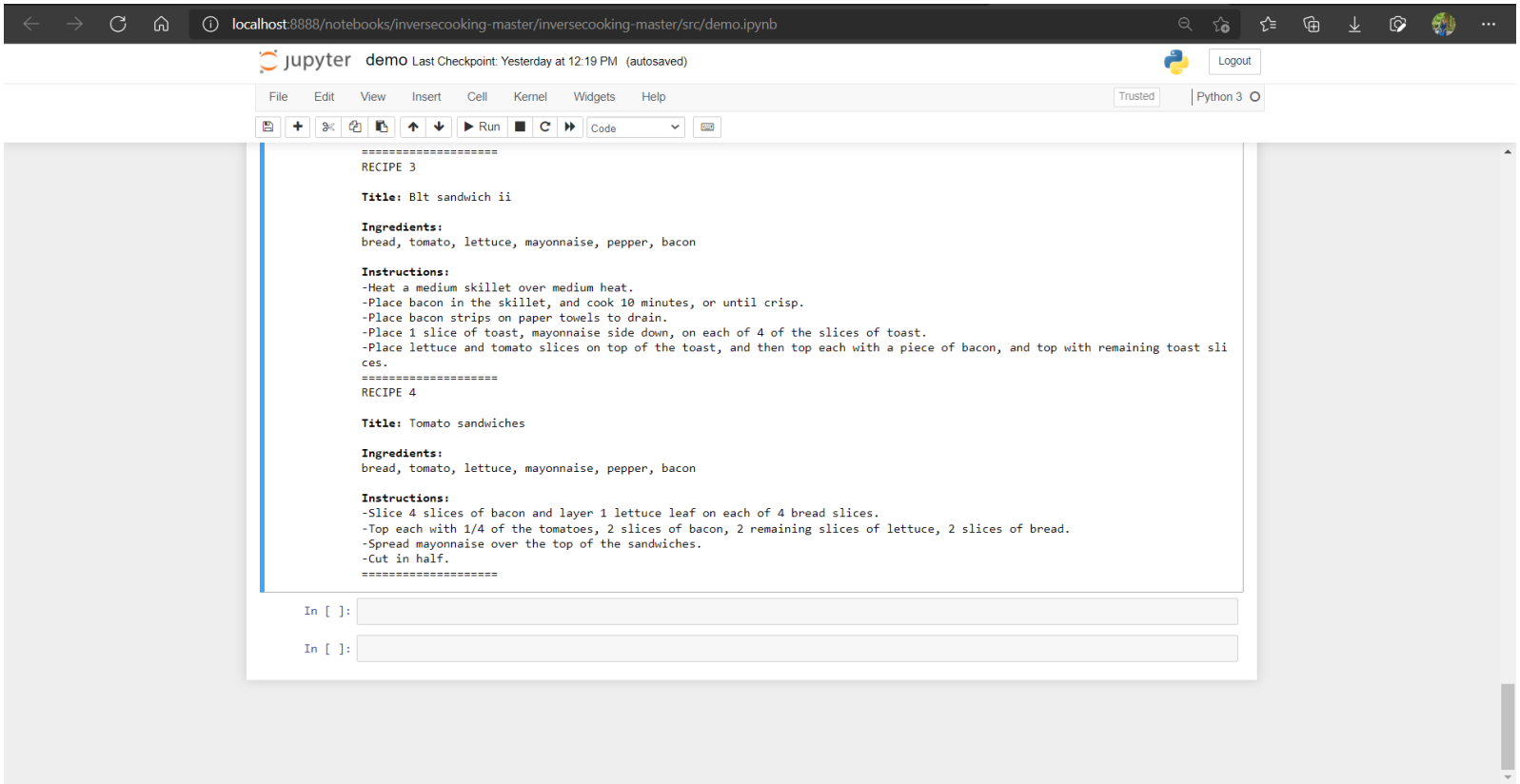
Sample Result Cont...



Sample Result Cont...



Sample Result Cont...



Conclusion

- The image-to-recipe generation system, which takes a food image and produces a recipe consisting of a title, ingredients and sequence of cooking instructions. We first predicted sets of ingredients from food images, showing that modeling dependencies matters. Then, we explored instruction generation conditioned on images and inferred ingredients, highlighting the importance of reasoning about both modalities at the same time.
- The accuracy of food detection can be made even more accurate by training the layers.

Reference I

1. Recipe1M+: A Dataset for Learning Cross-Modal Embeddings for Cooking Recipes and Food Images – MIT

2. Food Detection with Image Processing Using Convolutional Neural Network (CNN) Method — IEEE Conference Publication — IEEE Xplore

3. (PDF) Food Image Recognition by Using Convolutional Neural Networks (CNNs) (researchgate.net)

4. simplified-recipes-1M Dataset - Dominik Schmidt Slide 1 (nvidia.com)

5. [PDF] Deep Learning Based Food Recognition — Semantic Scholar

6. Food Ingredients and Recipes Dataset with Images — Kaggle

.

Thank You