

AIMS-DTU Research Intern Round 2

From Picture to Plate

Overview

This project, *From Picture to Plate*, focuses on leveraging Vision-Language Models (VLMs) to generate concise 2–3 step cooking instructions from food images and noisy or vague dish titles. The task emphasizes creativity in data curation, prompt engineering, and understanding the alignment between visual and textual data to produce practical recipe summaries.

Objective

The goal is to create a pipeline that takes a food image and a noisy or vague dish title as input and generates a concise 2–3 step cooking instruction that captures the essence of the dish. The pipeline should generalize to unseen image-title pairs using a pretrained VLM and few-shot prompting.

Dataset Setup

The dataset was curated from Food.com, a popular recipe website, to ensure diversity in dishes and reliable recipe data. The process involved:

- **Data Collection:** Scraped 15 food image samples along with their corresponding noisy titles (e.g., "cheesy bake", "noodly thing") and full cooking instructions from Food.com.
- **Manual Summary Creation:** For each of the 15 samples, concise 2–3 step cooking summaries were manually written based on the full instructions. These were stored as tuples in a Python dictionary and exported to a CSV file (`y_train.csv`) using the pandas library.
- **Test Dataset:** A separate test set of 10 image-title pairs was created similarly, with manually written summaries stored in `y_test.csv` for evaluation.
- **Image Handling:** Images were downloaded and stored locally, with paths

updated in `x_train_noisier.csv` and `x_test_noisier.csv` to point to the correct file locations (e.g., 1.png, 2.png, etc as sample id to map with the written summaries) with the most vague and noisy title to think of .

`x_test_noisier` with the noisy labels (like lemon baked cud as zingy baked fishy delight)

```

C: > Users > divan > Downloads > test_dataset > x_test_noisier.csv > data
1  sample_id,image_path,noisy_title
2  16,C:\Users\divan\Downloads\test_dataset\16.png,zingy baked fishy delight
3  17,C:\Users\divan\Downloads\test_dataset\17.png,saucy oven rib thing
4  18,C:\Users\divan\Downloads\test_dataset\18.png,juicy grill chunk bomb
5  19,C:\Users\divan\Downloads\test_dataset\19.png,creamy spicy noodly birdy thing
6  20,C:\Users\divan\Downloads\test_dataset\20.png,feathery chunky pie mess
7  21,C:\Users\divan\Downloads\test_dataset\21.png,yeasty beer loaf situation
8  22,C:\Users\divan\Downloads\test_dataset\22.png,crispy oven starchy chunks
9  23,C:\Users\divan\Downloads\test_dataset\23.png,cold tuber mayo deal
10 24,C:\Users\divan\Downloads\test_dataset\24.png,hot italian soupy stuff
11 25,C:\Users\divan\Downloads\test_dataset\25.png,boozy sweet cluck thing
12

```

Similarly , `x_train_noisier` is also formed (like pasta toss is replaced by very vague title easy noodly throw together)

```

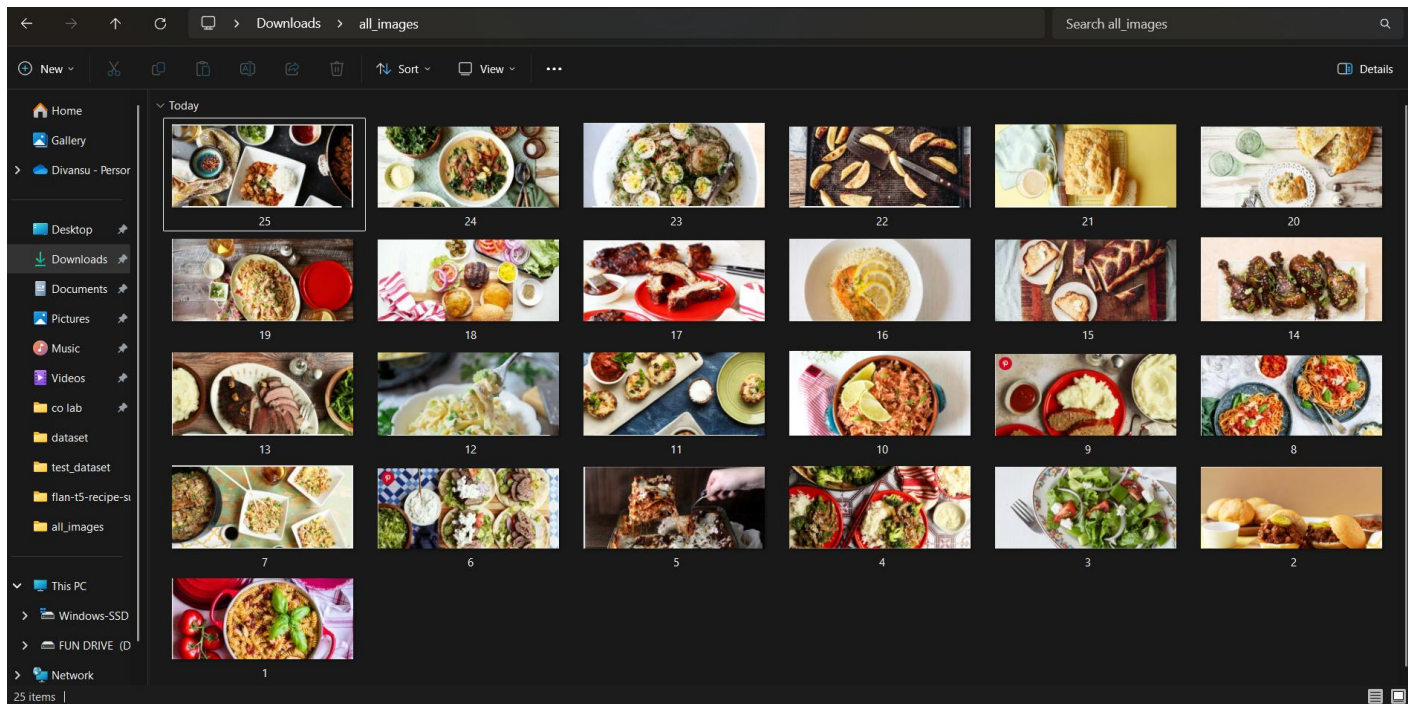
C: > Users > divan > Downloads > dataset > x_train_noisier.csv > data
1  sample_id,image_path,noisy_title
2  1,C:\Users\divan\Downloads\dataset\1.png,easy noodly throw together
3  2,C:\Users\divan\Downloads\dataset\2.png,messy meat bun deal
4  3,C:\Users\divan\Downloads\dataset\3.png,top-tier veggie feta chaos
5  4,C:\Users\divan\Downloads\dataset\4.png,lazy sizzling cow-broccoli thing
6  5,C:\Users\divan\Downloads\dataset\5.png,layered cheesy noodly brick
7  6,C:\Users\divan\Downloads\dataset\6.png,wrap of spinning moo strips
8  7,C:\Users\divan\Downloads\dataset\7.png,Col 2: image_path y ricey birdy wok mix
9  8,C:\Users\divan\Downloads\dataset\8.png,legendary noodles from jo mama
10 9,C:\Users\divan\Downloads\dataset\9.png,oven d gobbler loaf
11 10,C:\Users\divan\Downloads\dataset\10.png,spicy fiesta rice chaos
12 11,C:\Users\divan\Downloads\dataset\11.png,cheesy loaf bites of pizza dreams
13 12,C:\Users\divan\Downloads\dataset\12.png,pan-made fishy bake
14 13,C:\Users\divan\Downloads\dataset\13.png,ideal cow chunk roast-a-thon
15 14,C:\Users\divan\Downloads\dataset\14.png,mom-style soy birdy bites
16 15,C:\Users\divan\Downloads\dataset\15.png,fluffy golden twisty bread
17

```

All images in the zip file are as shown , (with 1 , 215 and the rest going to the testing dataset)

The training dataset included dishes like **mom-style soy birdy bites**, with summaries formatted as three-step instructions. For example:

- **Title:** mom-style soy birdy bites
- **Summary:**
 - "Step 1: Simmer all ingredients for 20 minutes, remove scum.
 - Step 2: Increase heat, reduce liquid to glaze, turning chicken.
 - Step 3: Arrange chicken, spoon glaze over, remove garlic and chili."



Task Flow

Step 1: Data Collection & Manual Summary Creation

The dataset was collected from Food.com, focusing on recipes with clear images and detailed instructions. For each of the 15 training samples, full instructions were condensed into 2–3 step summaries, ensuring clarity and brevity. These summaries served as ground truth for few-shot prompting. **The test set of 10 samples** was prepared similarly to evaluate model generalization.

Step 2: Model Selection & Pipeline Design

Initially, a Facebook VLM was tested but proved unsuitable, as it generated irrelevant outputs not aligned with the task of summarizing recipes. After researching suitable models, **Google’s FLAN-T5-Large** was selected due to its **strong performance in text generation and ability to handle structured prompts effectively**. The pipeline was designed as follows:

- **Tokenizer and Model:** Used `AutoTokenizer` and `AutoModelForSeq2SeqLM` from HuggingFace's Transformers library to process inputs and generate outputs.
- **Prompt Engineering:** A few-shot prompting approach was adopted, where three randomly selected training examples (title-summary pairs) were included in the prompt to guide the model. The prompt format was:
Generate a concise 2–3 step recipe for the given dish title.
Title: [Example Title 1]
Recipe: [Example Summary 1]
...
Title: [Input Title]
Recipe:
- **Inference Settings:** The model used temperature (0.7), top-k sampling (50), and top-p sampling (0.95) to balance creativity and coherence in the generated summaries.
- **Image Verification:** Images were loaded using the PIL library to confirm accessibility and correct format (RGB), though the current pipeline focused on text-based inputs (titles) for simplicity.

The pipeline was implemented in Python, utilizing libraries like `transformers`, `torch`, `pandas`, `tqdm`, and `zipfile`. A GPU (NVIDIA GeForce RTX 3050 with CUDA 12.4) was used to accelerate inference.

Step 3: Inference on Test Set

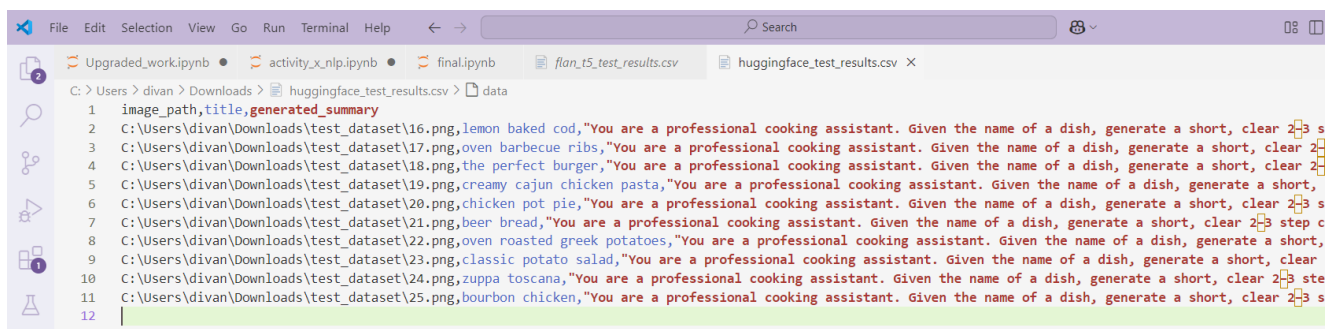
The pipeline was tested on the 10 test samples, generating summaries for each noisy title. *The outputs were saved to `flan_t5_test_results.csv`. Example outputs include:*

- **Title:** zingy baked fishy delight
- **Generated Summary:**
 - Step 1: Preheat oven to 450°F
 - Step 2: Toss the tuna with the lemon zest and remaining ingredients
 - Step 3: Bake until fish is flaky and whites are opaque, about 25 minutes"
- **Title:** creamy spicy noodly birdy thing
- **Generated Summary:**
 - Step 1: Scoop up the cream cheese and onion mixture
 - Step 2: Add the chopped jalapeno and the chopped parsley
 - Step 3: Spread the cheese mixture on the patty."

Step 4: Evaluation and Reflection

The generated summaries were manually compared to the original test summaries in `y_test.csv`. Key observations include:

- **Successes:** The final model correctly captured the essence of dishes in most cases, producing concise and actionable instructions. For example, the summary for *"juicy grill chunk bomb"*,
 - Step 1: Preheat the grill to a medium-high heat
 - Step 2: Place the chicken in a large skillet and brush with olive oil
 - Step 3: Add the chunks of garlic, parsley, and parsley"
- **Failures:** In some cases, the model oversimplified complex dishes or missed specific ingredients (e.g., omitting garnishes like parsley), likely due to the limited number of few-shot examples. A significant challenge was integrating the PC's GPU (*NVIDIA GeForce RTX 3050 with CUDA 12.4*). Without proper GPU configuration, the model would freeze during inference, causing the VS Code editor to crash. This was overcome by ensuring compatibility with PyTorch and CUDA, enabling stable and faster processing.
- **Common Failure Patterns:** The model occasionally produced generic steps (e.g., "cook and serve") when the noisy title was too vague, underscoring the need for descriptive prompts. Early attempts with the **BLIP model failed**, as it merely echoed the few-shot prompt examples without generating meaningful summaries, indicating it was not suited for this task. Similarly, **the facebook/bart-large-cnn model produced irrelevant outputs, such as repeating the prompt itself** (e.g., for every title, it generated: "You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."), confirming its inadequacy as a summarizer for this context.
- **Blip model output :- (only echoing my prompt) :**



```
1 image_path,title,generated_summary
2 C:\Users\divan\Downloads\test_dataset\16.png,lemon baked cod,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
3 C:\Users\divan\Downloads\test_dataset\17.png,oven barbecue ribs,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
4 C:\Users\divan\Downloads\test_dataset\18.png,the perfect burger,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
5 C:\Users\divan\Downloads\test_dataset\19.png,creamy cajun chicken pasta,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
6 C:\Users\divan\Downloads\test_dataset\20.png,chicken pot pie,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
7 C:\Users\divan\Downloads\test_dataset\21.png,beer bread,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
8 C:\Users\divan\Downloads\test_dataset\22.png,oven roasted greek potatoes,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
9 C:\Users\divan\Downloads\test_dataset\23.png,classic potato salad,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
10 C:\Users\divan\Downloads\test_dataset\24.png,zuppa toscana,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
11 C:\Users\divan\Downloads\test_dataset\25.png,bourbon chicken,"You are a professional cooking assistant. Given the name of a dish, generate a short, clear 2-3 step cooking recipe..."
12
```

facebook/bart-large-cnn model output (irrelevant output line repeating

Step 2: Add your own creative cooking touch.

Step 3: Add your own creative cooking touch.” In many of the outputs)

```
File Edit Selection View Go Run Terminal Help VLM for food ingredients
ealry.ipynb bark_earlier.csv flan_t5_test_results.csv final.ipynb x_train_fixed.csv x_train_noisier.csv x_test_noisier.csv
C:\Users\divan> Downloads\flan_t5_results> facebook> bark_earlier.csv> data
1 sample_id,image_path,title,generated_summary
2 16,C:\Users\divan\Downloads\test_dataset\16.png,zingy baked fishy delight,"Step 1: 1: Cook pasta in boiling water; drain.
3 Step 2: 2: In a large bowl, stir in the cream cheese, Parmesan, onion, tomatoes, and salt.
4 Step 3: 3: Add the penne, and mix until combined."
5 17,C:\Users\divan\Downloads\test_dataset\17.png,saucy oven rib thing,"Step 1: 1: Preheat the oven to 450 degrees F. Prepare the oven ribs. Place the ribs in a large
6 Step 2: Add your own creative cooking touch.
7 Step 3: Add your own creative cooking touch."
8 18,C:\Users\divan\Downloads\test_dataset\18.png,juicy grill chunk bomb,"Step 1: 1: Preheat grill to medium-high heat. Add onion, garlic, and pepper. Cook, stirring,
9 Step 2: Add your own creative cooking touch.
10 Step 3: Add your own creative cooking touch."
11 19,C:\Users\divan\Downloads\test_dataset\19.png,creamy spicy noodly birdy thing,"Step 1: 1: 12 cup cream cheese 1 tablespoon chopped jalapeno 1 tablespoon chopped
12 Step 2: Add your own creative cooking touch.
13 Step 3: Add your own creative cooking touch."
14 20,C:\Users\divan\Downloads\test_dataset\20.png,feathery chunky pie mess,"Step 1: 1: Melt butter and add sugar.
15 Step 2: 2: Heat up a large skillet and add flour.
16 Step 3: 3: Add eggs and stir."
17 21,C:\Users\divan\Downloads\test_dataset\21.png,yeasty beer loaf situation,"Step 1: 1: In a large bowl, combine the yeast and sugar.
18 Step 2: 2: Add the flour, salt, and beer.
19 Step 3: 3: Pour the mixture into a loaf pan, and bake for about 45 minutes at 375°F."
20 22,C:\Users\divan\Downloads\test_dataset\22.png,crispy oven starchy chunks,"Step 1: 1: Preheat oven to 350°F (175°C). Line an 8x8x2-inch baking pan with foil.
21 Step 2: 2: Place shredded potatoes and cauliflower in the preheated oven.
22 Step 3: 3: Bake, tossing frequently, until potatoes are crisp and cauliflower is cooked through, about 15 minutes."
23 23,C:\Users\divan\Downloads\test_dataset\23.png,cold tuber mayo deal,"Step 1: 1: Cut 2 ounces tubers lengthwise in half. Cut each half in half. Cut each half in half
24 Step 2: Add your own creative cooking touch.
25 Step 3: Add your own creative cooking touch."
26 24,C:\Users\divan\Downloads\test_dataset\24.png,hot italian soupy stuff,"Step 1: 1: Add 1 pound diced tomatoes to a large bowl.
27 Step 2: 2: Add 2 pounds diced peppers.
28 Step 3: 3: Stir to mix."
29 25,C:\Users\divan\Downloads\test_dataset\25.png,boozy sweet cluck thing,"Step 1: 1: Make boozy sweet cluck thing.
30 Step 2: 2: Prepare the ingredients.
31 Step 3: 3: Pour the ingredients into a bowl."
32
```

1) Evaluation :

Y_test is provided (actual summaries taken from food.com to manually match the summaries) :-

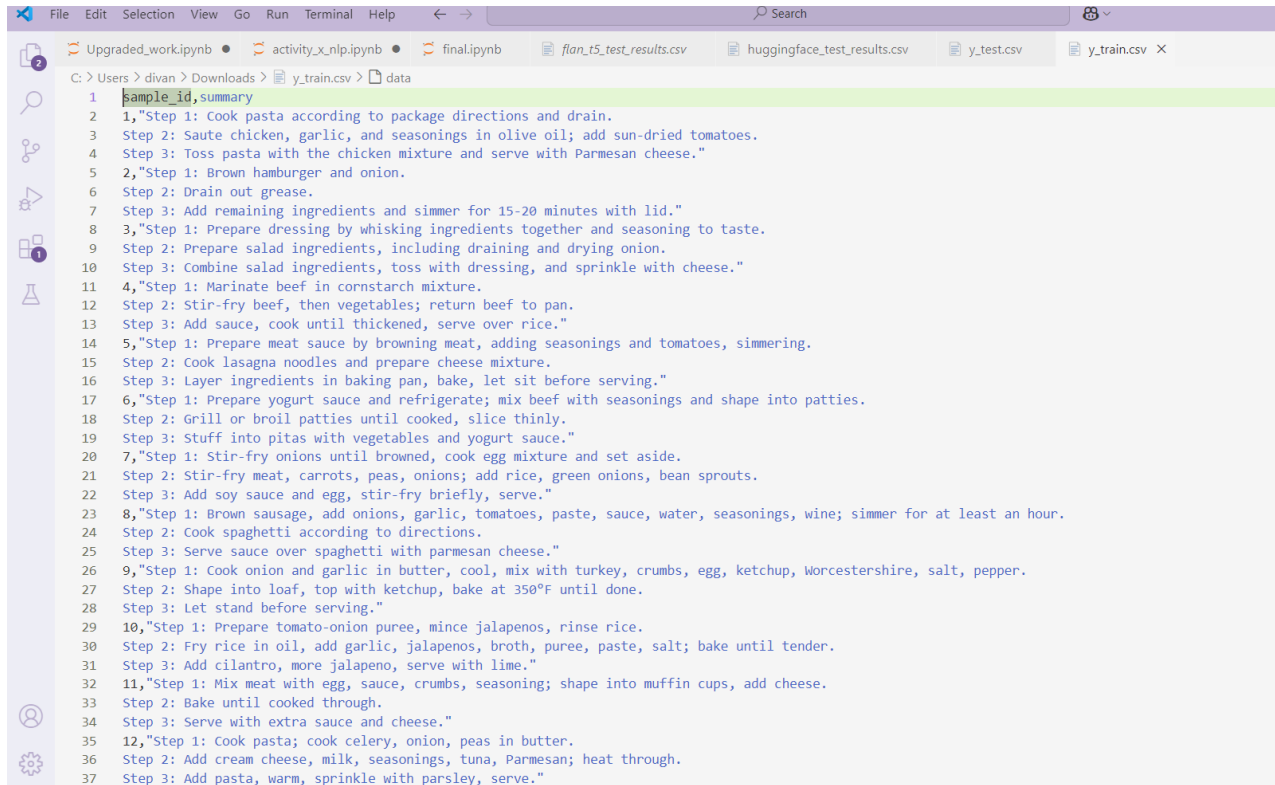
```
File Edit Selection View Go Run Terminal Help Search
Upgraded_work.ipynb activity_x_nlp.ipynb final.ipynb flan_t5_test_results.csv huggingface_test_results.csv y_test.csv
C:\Users\divan> Downloads> y_test.csv> data
1 sample_id,summary
2 16,"Step 1: Mix butter and lemon juice; coat fish with flour, salt, and pepper mixture.
3 Step 2: Place fish in a baking dish, pour remaining butter mixture over, sprinkle with paprika.
4 Step 3: Bake at 350°F for 25-30 minutes, garnish with parsley and lemon slices."
5 17,"Step 1: Season ribs with salt, pepper, and liquid smoke; apply brown sugar, paprika, garlic powder, and red pepper rub.
6 Step 2: Bake ribs in foil at 300°F for 2-2.5 hours until tender.
7 Step 3: Cut ribs, brush with sauce, broil 1-2 minutes per side until bubbly."
8 18,"Step 1: Mix egg, seasonings, and beef; gently shape into patties with a thumb depression.
9 Step 2: Grill burgers at medium heat for 6-8 minutes per side until no longer pink (160°F).
10 Step 3: Serve in a bun with your favorite toppings."
11 19,"Step 1: Toss chicken with Cajun seasoning; sauté in butter until tender.
12 Step 2: Add green onion, cream, tomatoes, basil, salt, garlic powder, and pepper; heat through.
13 Step 3: Serve over hot linguine, toss with Parmesan cheese."
14 20,"Step 1: Sauté onion, celery, carrots, and potatoes in margarine; add flour and cook 1 minute.
15 Step 2: Stir in broth, half and half, salt, pepper, and chicken; cook until thickened.
16 Step 3: Pour into a dish, top with pie shell, bake at 400°F for 40-50 minutes."
17 21,"Step 1: Mix dry ingredients and beer.
18 Step 2: Pour into a greased loaf pan, top with melted butter.
19 Step 3: Bake at 375°F for 1 hour, cool for 15 minutes."
20 22,"Step 1: Toss potatoes with all ingredients, season with salt and pepper.
21 Step 2: Bake at 420°F for 40 minutes until golden-brown.
22 Step 3: Stir, season again, add water if needed, bake another 40 minutes."
23 23,"Step 1: Boil potatoes, cool, and dice.
24 Step 2: Mix with mayonnaise, vinegar, sugar, mustard, salt, garlic powder, pepper, celery, onions, and eggs.
25 Step 3: Sprinkle with paprika and serve."
26 24,"Step 1: Cook sausage with onions, garlic, salt, pepper, and red pepper flakes.
27 Step 2: Add potatoes, stock, and water; cook until potatoes are done, then add flour and simmer.
28 Step 3: Stir in kale, bacon, and cream; serve."
29 25,"Step 1: Brown chicken in oil, remove from skillet.
30 Step 2: Add remaining ingredients, bring to a boil, then simmer with chicken for 20 minutes.
31 Step 3: Serve over hot rice."
32
```


- **Model Choice Justification:** After extensive research across platforms to identify a VLM capable of capturing recipe complexity and generalizing well in the local environment, **Google's FLAN-T5-Large was selected**. Unlike the **BLIP model**, which failed to generate meaningful summaries and just echoed the few shots prompts, and the **facebook/bart-large-cnn** model, which produced irrelevant outputs, FLAN-T5-Large demonstrated robust text generation and accurate generalization from few-shot examples without requiring fine-tuning. Its compatibility with the GPU setup and ability to handle structured prompts made it ideal for this task.
- **Design Decisions:** The decision to focus on text-based inputs (titles) rather than images simplified the pipeline, as image processing required additional preprocessing. Few-shot prompting was effective given the small dataset size, avoiding the need for resource-intensive fine-tuning.

Deliverables

1. **Notebook/Script:** A Python script was developed to load data, process images, generate summaries using FLAN-T5, and save results. The script includes clear console logging for debugging and progress tracking.
2. **Manually Written Summaries:** 15 training and 10 test summaries were created and stored in **y_train.csv** and **y_test.csv**, respectively.

Attached y_train here for reference : -



```

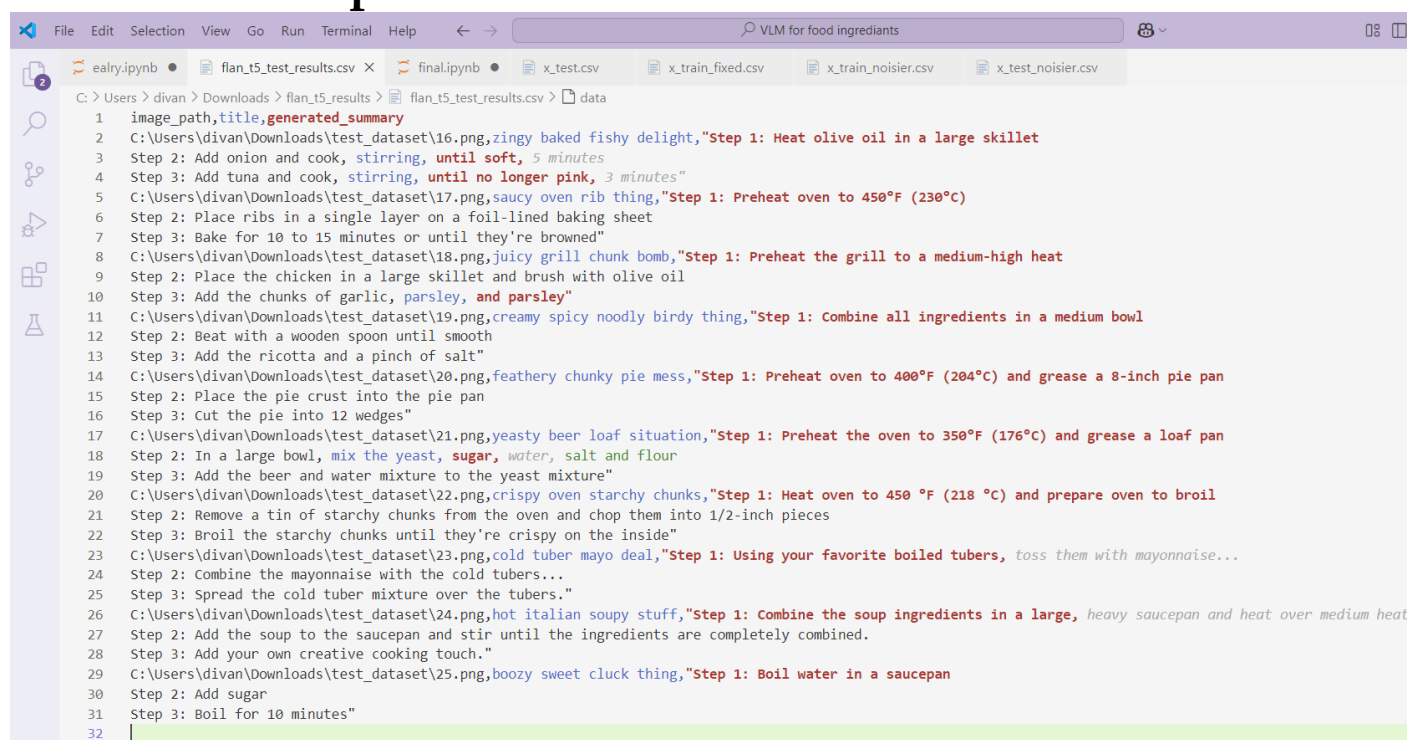
1 sample_id,summary
2 1,"Step 1: Cook pasta according to package directions and drain.
3 Step 2: Saute chicken, garlic, and seasonings in olive oil; add sun-dried tomatoes.
4 Step 3: Toss pasta with the chicken mixture and serve with Parmesan cheese."
5 2,"Step 1: Brown hamburger and onion.
6 Step 2: Drain out grease.
7 Step 3: Add remaining ingredients and simmer for 15-20 minutes with lid."
8 3,"Step 1: Prepare dressing by whisking ingredients together and seasoning to taste.
9 Step 2: Prepare salad ingredients, including draining and drying onion.
10 Step 3: Combine salad ingredients, toss with dressing, and sprinkle with cheese."
11 4,"Step 1: Marinate beef in cornstarch mixture.
12 Step 2: Stir-fry beef, then vegetables; return beef to pan.
13 Step 3: Add sauce, cook until thickened, serve over rice."
14 5,"Step 1: Prepare meat sauce by browning meat, adding seasonings and tomatoes, simmering.
15 Step 2: Cook lasagna noodles and prepare cheese mixture.
16 Step 3: Layer ingredients in baking pan, bake, let sit before serving."
17 6,"Step 1: Prepare yogurt sauce and refrigerate; mix beef with seasonings and shape into patties.
18 Step 2: Grill or broil patties until cooked, slice thinly.
19 Step 3: Stuff into pitas with vegetables and yogurt sauce."
20 7,"Step 1: Stir-fry onions until browned, cook egg mixture and set aside.
21 Step 2: Stir-fry meat, carrots, peas, onions; add rice, green onions, bean sprouts.
22 Step 3: Add soy sauce and egg, stir-fry briefly, serve."
23 8,"Step 1: Brown sausage, add onions, garlic, tomatoes, paste, sauce, water, seasonings, wine; simmer for at least an hour.
24 Step 2: Cook spaghetti according to directions.
25 Step 3: Serve sauce over spaghetti with parmesan cheese."
26 9,"Step 1: Cook onion and garlic in butter, cool, mix with turkey, crumbs, egg, ketchup, Worcestershire, salt, pepper.
27 Step 2: Shape into loaf, top with ketchup, bake at 350°F until done.
28 Step 3: Let stand before serving."
29 10,"Step 1: Prepare tomato-onion puree, mince jalapenos, rinse rice.
30 Step 2: Fry rice in oil, add garlic, jalapenos, broth, puree, paste, salt; bake until tender.
31 Step 3: Add cilantro, more jalapeno, serve with lime."
32 11,"Step 1: Mix meat with egg, sauce, crumbs, seasoning; shape into muffin cups, add cheese.
33 Step 2: Bake until cooked through.
34 Step 3: Serve with extra sauce and cheese."
35 12,"Step 1: Cook pasta; cook celery, onion, peas in butter.
36 Step 2: Add cream cheese, milk, seasonings, tuna, Parmesan; heat through.
37 Step 3: Add pasta, warm, sprinkle with parsley, serve."

```

3. **Model-Generated Outputs:** Summaries for the test set were saved in `flan_t5_test_results.csv`.
4. **Video Walkthrough:** A 5 minute video is attached to explain the pipeline, prompt structure, model selection reasoning, and two sample outputs (e.g., zingy Baked fishy delight and juicy grill chunk bomb).

Conclusion

Generated Output



The screenshot shows a Jupyter Notebook interface with a file explorer at the top displaying files like `flan_t5_test_results.csv` and `flan_t5_test_results.csv`. The main area shows a code cell with a table of generated cooking instructions. The table has three columns: `image_path`, `title`, and `generated_summary`. The data rows show various dishes like 'zingy baked fishy delight', 'saucy oven rib thing', 'juicy grill chunk bomb', 'creamy spicy noodly birdy thing', 'feathery chunky pie mess', 'yeasty beer loaf situation', 'crispy oven starchy chunks', 'cold tuber mayo deal', and 'hot italian soupy stuff', each followed by a detailed step-by-step cooking instruction.

```

1  image_path,title,generated_summary
2  C:\Users\divan\Downloads\test_dataset\16.png,zingy baked fishy delight,"Step 1: Heat olive oil in a large skillet
3  Step 2: Add onion and cook, stirring, until soft, 5 minutes
4  Step 3: Add tuna and cook, stirring, until no longer pink, 3 minutes"
5  C:\Users\divan\Downloads\test_dataset\17.png,saucy oven rib thing,"Step 1: Preheat oven to 450°F (230°C)
6  Step 2: Place ribs in a single layer on a foil-lined baking sheet
7  Step 3: Bake for 10 to 15 minutes or until they're browned"
8  C:\Users\divan\Downloads\test_dataset\18.png,juicy grill chunk bomb,"Step 1: Preheat the grill to a medium-high heat
9  Step 2: Place the chicken in a large skillet and brush with olive oil
10 Step 3: Add the chunks of garlic, parsley, and parsley"
11 C:\Users\divan\Downloads\test_dataset\19.png,creamy spicy noodly birdy thing,"Step 1: Combine all ingredients in a medium bowl
12 Step 2: Beat with a wooden spoon until smooth
13 Step 3: Add the ricotta and a pinch of salt"
14 C:\Users\divan\Downloads\test_dataset\20.png,feathery chunky pie mess,"Step 1: Preheat oven to 400°F (204°C) and grease a 8-inch pie pan
15 Step 2: Place the pie crust into the pie pan
16 Step 3: Cut the pie into 12 wedges"
17 C:\Users\divan\Downloads\test_dataset\21.png,yeasty beer loaf situation,"Step 1: Preheat the oven to 350°F (176°C) and grease a loaf pan
18 Step 2: In a large bowl, mix the yeast, sugar, water, salt and flour
19 Step 3: Add the beer and water mixture to the yeast mixture"
20 C:\Users\divan\Downloads\test_dataset\22.png,crispy oven starchy chunks,"Step 1: Heat oven to 450 °F (218 °C) and prepare oven to broil
21 Step 2: Remove a tin of starchy chunks from the oven and chop them into 1/2-inch pieces
22 Step 3: Broil the starchy chunks until they're crispy on the inside"
23 C:\Users\divan\Downloads\test_dataset\23.png,cold tuber mayo deal,"Step 1: Using your favorite boiled tubers, toss them with mayonnaise...
24 Step 2: Combine the mayonnaise with the cold tubers...
25 Step 3: Spread the cold tuber mixture over the tubers."
26 C:\Users\divan\Downloads\test_dataset\24.png,hot italian soupy stuff,"Step 1: Combine the soup ingredients in a large, heavy saucepan and heat over medium heat
27 Step 2: Add the soup to the saucepan and stir until the ingredients are completely combined.
28 Step 3: Add your own creative cooking touch."
29 C:\Users\divan\Downloads\test_dataset\25.png,boozy sweet cluck thing,"Step 1: Boil water in a saucepan
30 Step 2: Add sugar
31 Step 3: Boil for 10 minutes"
32

```

The FLAN-T5-Large model, combined with few-shot prompting (randomly selecting 3 examples from training set), successfully generated concise and practical cooking instructions from noisy dish titles. The pipeline demonstrated the potential of pretrained VLMs for recipe summarization, even without fine-tuning. Future improvements could include incorporating image data for multimodal processing or expanding the dataset for better generalization.

Divansu Mishra (24/A03/021)