

# Inverse Cooking: Recipe Generation from Food Images

## (Supplementary Material)

This supplementary material intends to provide further details as well as qualitative results. In Section 1, we describe additional implementation and training details. Section 2 presents an analysis of our ingredient vocabulary before and after its pre-processing. Examples of generated recipes, displayed together with real ones from the dataset, are presented in Section 3. Section 4 includes screenshots of the two forms that were used to collect data for the user studies. Section 5 includes examples of human written recipes compared to real and generated ones. Finally, in Section 5.1, we provide examples of generated recipes for out-of-dataset pictures taken by authors.

### 1. Training Details

**Ingredient Prediction.** Feed-forward models  $FF_{BCE}$ ,  $FF_{TD}$  and  $FF_{IOU}$  were trained with a mini-batch size of 300, whereas  $FF_{DC}$  was trained with a mini-batch size of 256. All of them were trained with a learning rate of 0.001. The learning rate for pre-trained ResNet layers was scaled for each model as follows:  $0.01 \times$  for  $FF_{BCE}$ ,  $FF_{IOU}$  and  $FF_{DC}$  and  $0.1 \times$  for  $FF_{TD}$ . Transformer list-based models  $TF_{list}$  were trained with mini-batch size 300 and learning rate 0.001, scaling the learning rate of ResNet layers with a factor of  $0.1 \times$ . Similarly, the set transformer  $TF_{set}$  was trained with mini-batch size of 300 and a learning rate of 0.0001, scaling the learning rate of pre-trained ResNet layers with a factor of  $1.0 \times$ . The optimization of  $TF_{set}$  minimizes a cost function composed of three terms, namely the ingredient prediction loss  $\mathcal{L}_{ingr}$  and the end-of-sequence loss  $\mathcal{L}_{eos}$  and the cardinality penalty  $\mathcal{L}_{card}$ . We set the contribution of each term with weights 1000.0 and 1.0 and 1.0, respectively. We use a label smoothing factor of 0.1 for all models trained with BCE loss ( $FF_{BCE}$ ,  $FF_{DC}$ ,  $TF_{set}$ ), which we found experimentally useful.

**Instruction Generation.** We use a batch size of 256 and learning rate of 0.001. Parameters of the image encoder module are taken from the ingredient prediction model and frozen during training for instruction generation.

All models are trained with Adam optimizer ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.99$  and  $\epsilon = 1e-8$ ), exponential decay of 0.99 after each epoch, dropout probability 0.3 and a maximum number of 400 epochs (if early stopping criterion is not met).

During training we randomly flip ( $p = 0.5$ ), rotate ( $\pm 10$  degrees) and translate images ( $\pm 10\%$  image size on each axis) for augmentation.

### 2. Ingredient Analysis

We provide visualizations of the ingredient vocabulary used to train our models. Figure 1 displays each unique ingredient in the vocabulary before and after our pre-processing stage. The size of each ingredient word indicates its frequency in the dataset (e.g. *butter* and *salt* appear in many recipes). After filtering and clustering ingredients, the distribution slightly changes (e.g. *pepper* becomes the most frequent ingredient, and popular ingredients such as *olive oil* or *vegetable oil* are clustered into *oil*). Additionally, we illustrate the high ingredient overlap in the dataset with an example of the different types of *cheese* that appear as different ingredients before pre-processing.

### 3. Generated Recipes

Figure 2 shows additional examples of generated recipes obtained with our method. We also provide the real recipe for completeness. Although sometimes far from the real recipe, our system is able to generate plausible and structured recipes for the input images. Common mistakes include failures in ingredient recognition (e.g. *stuffed tomatoes* are confused with *stuffed peppers* in Figure 2b), inconsistencies between ingredients and instructions (e.g. *cucumber* is predicted as an ingredient but unused in Figure 2d, and *meat* is mentioned in the title and instructions but is not predicted as an ingredient in Figure 2e), and repetitions in ingredient enumeration (e.g. *Stir in tomato sauce, tomato paste, tomato paste, ...* in Figure 2c).

### 4. User Study Forms

We provide screenshots of the two forms used to collect data for user studies. Figure 3 shows the interface used by users to select image ingredients (each ingredient was selected using a drop-down menu), and write recipes (as free-form text). Figure 4 shows the form we used to assess whether a recipe matched the provided image according to human judgment.

## 5. Human-written Recipes

In Figure 5 we show examples of recipes written by humans, which were collected using the form in Figure 3. We also display the real and generated recipes for completeness. Recipes written by humans tend to be shorter, with an average of 5.29 instructions of 9.03 words each. In contrast, our model generates recipes that contain an average of instructions 9.21 of 9 words each, which closely matches the real distribution (9.08 sentences of length 12.79).

### 5.1. Dine Out: A case study

We test the capabilities of our model to generalize for out-of-dataset food images. Figure 6 shows recipes obtained for food images taken by authors at their homes or in restaurants during the weeks prior to the submission.



Figure 2: **Recipe examples.** We show both real and generated recipes for different test images.



GENERATED RECIPE

**Title:** Sloe gin martini

**Ingredients:**

juice, vodka, orange, ice, liqueur, lemon, gin

**Instructions:**

- In a cocktail shaker, combine vodka, gin, cointreau, cranberry juice, and lemon juice.
- Cover and shake well.
- Strain into a chilled martini glass.
- Garnish with orange slice.

---

REAL RECIPE

**Title:** Vampire kiss martini

**Ingredients:**

juice, vodka, lemon

**Instructions:**

- In a shaker filled with ice, combine vodka and juice.
- Strain into martini glass and garnish with lemon twist.





#### GENERATED RECIPE

**Title:** Stuffed peppers

**Ingredients:**

pepper, tomato, onion, spinach, clove, salt, cheese, mushroom, oil, rice, water

**Instructions:**

- Preheat oven to 350 degrees f (175 degrees c).
- Heat olive oil in a skillet over medium heat.
- Cook and stir onion and garlic in hot oil until onion is translucent, 5 to 7 minutes.
- Stir rice, water, salt, and black pepper into onion mixture; bring to a boil.
- Reduce heat to low, cover skillet, and simmer until rice is tender and water is absorbed, about 20 minutes.
- Stir spinach into mushroom mixture; cook and stir until spinach is wilted, about 5 minutes.
- Remove skillet from heat.
- Stir ricotta cheese into mushroom mixture; spoon into peppers.
- Place stuffed peppers into a baking dish.
- Bake in the preheated oven until peppers are tender, about 30 minutes

---

#### REAL RECIPE

**Title:** Vera cruz tomatoes

**Ingredients:**

tomato, bacon, onion, spinach, cream, pepper, salt, cheese

**Instructions:**

- Preheat an oven to 375 degrees f (190 degrees c).
- Grease an 8x8 inch baking dish.
- Cut tops from tomatoes; remove seeds and membranes.
- Place tomato shells upside down on paper towels to drain until filling.
- Place the bacon in a large, deep skillet, and cook over medium-high heat, turning occasionally, until evenly browned, about 10 minutes.
- Drain the bacon slices on a paper towel-lined plate, reserve drippings.
- Crumble bacon and set aside.
- Return 2 tablespoons of drippings to skillet.
- Cook onion in the bacon drippings until tender, about 5 minutes.
- Stir in spinach; cook and stir until wilted, about 2 minutes.
- Remove from heat.
- Stir in sour cream, reserved crumbled bacon, and hot pepper sauce.
- Sprinkle tomato shells with salt; fill evenly with bacon and spinach mixture.
- Place tomatoes in prepared baking dish.
- Bake in preheated oven until hot, 20 to 25 minutes.
- Remove, and top evenly with the shredded cheese.
- Return to the oven; bake until cheese is melted, about 5 minutes.



#### GENERATED RECIPE

**Title:** Pasta with ground beef and sauce

**Ingredients:**

cheese, beef, onion, pasta, tomato, pepper, salt, garlic, clove, basil, oregano

**Instructions:**

- In a large skillet, brown ground beef, onion, garlic, and green pepper until meat is no longer pink; drain.
- Stir in tomato sauce, tomato paste, tomato paste, basil, oregano, salt, pepper, and garlic powder.
- Bring to a boil; reduce heat and simmer for 15 minutes.
- Meanwhile, cook pasta according to package directions.
- Drain.
- Stir in sauce and parmesan cheese.
- Serve over pasta.

---

#### REAL RECIPE

**Title:** Baked ziti

**Ingredients:**

pasta, oil, onion, clove, zucchini, mushroom, cheese, cream

**Instructions:**

- Cook pasta according to package directions for al dente.
- Then drain it and set aside.
- While pasta is cooking, heat olive oil in a large saute pan over medium high heat.
- Then add onion, garlic, zucchini and mushrooms and saute for about 10 minutes or until vegetables are tender.
- Add spaghetti sauce and simmer for 5 minutes.
- Preheat oven to 350 f. spray a large, deep baking dish (see note below) with cooking spray.
- Layer ingredients as follows: 1/3 of the sauce mixture, 1/2 of the cooked ziti, all of the provolone cheese, all of the sour cream, 1/2 of the remaining sauce mixture, remaining ziti, all of the mozzarella cheese, and the remaining sauce mixture.
- Bake for 30 minutes.
- Notes: this dish needs a very large and deep baking dish to accommodate all of the ingredients.
- I use a corningware 2.8 liter casserole dish, which is the equivalent to 2 1/2 quarts, and i usually have sauce left over.
- If you have a 3 quart baking dish, that would be the best size to use.
- Also, you want to make sure that all of the noodles are covered with sauce on the final layer so that they don't dry out when baking.



#### GENERATED RECIPE

**Title:** Antipasto pasta salad

**Ingredients:**

olive, pasta, cheese, pepper, onion, salami, artichoke, tomato, cucumber, salt, mayonnaise, vinegar, mustard

**Instructions:**

- Cook pasta according to package directions.
  - Drain and rinse with cold water.
  - In a large bowl, combine pasta, salami, artichoke hearts, red onion, olives, tomatoes, and cheese.
  - In a small bowl, whisk together mayonnaise, vinegar, mustard, salt and pepper.
  - Pour dressing over pasta mixture and toss to coat.
  - Cover and refrigerate for at least 1 hour before serving.
- 

#### REAL RECIPE

**Title:** Pesto pasta and chickpea salad

**Ingredients:**

pasta, chickpeas, italian\_dressing, pesto\_sauce, pimiento, olive, cheese

**Instructions:**

- Prepare bow-tie pasta according to package directions.
- Combine all ingredients in a large bowl and stir to mix well.
- Increasing or decreasing the italian dressing and pesto to taste.
- Serve at room temperature for the best flavor.



#### GENERATED RECIPE

**Title:** Asian beef stir-fry

**Ingredients:**

pepper, peas, sesame, clove, cilantro, pasta, oil, mushroom, linguine

**Instructions:**

- Cook pasta as directed on package, omitting salt.
- Meanwhile, heat oil in large nonstick skillet on medium-high heat.
- Add meat; stir-fry 5 min.
- Or until done.
- Add peppers, mushrooms and garlic; stir-fry 2 min.
- Add pasta sauce; stir-fry 2 min.
- Or until heated through.
- Drain pasta; place in large bowl.
- Add meat mixture; toss to coat.
- Sprinkle with cilantro.

---

#### REAL RECIPE

**Title:** Beef & noodles with fresh vegetables

**Ingredients:**

italian\_dressing, soy\_sauce, clove, peas, cilantro, spaghetti

**Instructions:**

- Mix dressing and soy sauce until blended.
- Stir-fry meat and garlic in 1 tbsp.
- Dressing mixture in large nonstick skillet on medium-high heat 3 min.
- Or until meat is barely pink in center.
- Transfer to plate; set aside.
- Add peas, peppers and remaining dressing mixture to skillet; stir-fry 3 min.
- Or until vegetables are crisp-tender.
- Return meat to skillet; stir-fry 1 min.
- Remove from heat; stir in tomatoes and cilantro.
- Serve over spaghetti.





#### GENERATED RECIPE

**Title:** Avocado soup

**Ingredients:**

pepper, salt, avocado, yogurt, clove, onion, cream, parsley, water

**Instructions:**

- In a large pot, combine the avocado, onion, garlic, parsley, salt, pepper, and water.
- Bring to a boil, reduce heat and simmer for 15 minutes.
- Remove from heat and let cool.
- In a blender, puree the avocado mixture until smooth.
- Stir in yogurt and sour cream.
- Serve immediately.

---

#### REAL RECIPE

**Title:** Basil pesto

**Ingredients:**

basil, nuts, clove, salt, pepper, oil, cheese

**Instructions:**

- In a blender, pulse the basil, pine nuts, garlic, salt, and pepper until finely chopped.
- With the blender still running, gradually add enough oil to form a smooth and thick consistency.
- Transfer the pesto to a medium bowl and stir in cheese.
- Season the pesto with more salt and pepper, to taste.
- The pesto can be made 2 days ahead.

Figure 3: **User Study 1.** Interface for writing recipes and selecting ingredients.

## QUALITATIVE EVALUATION FOR FOOD UNDERSTANDING METHODS

Predicting ingredients and writing recipes directly from food images are challenging tasks which require strong visual recognition and reasoning models to understand food, its components and preparation instructions.

We have designed several methods to automatically solve these tasks, which we need to evaluate. In order to do so, we need your cooperation (and we are very grateful for it!).

In this test, **you will be prompted with 3 different food images, and for each of them we ask you to:**

**a) choose at least 1 and up to 20 different ingredients that are needed to prepare what you see in the image. Note that you can leave empty slots. Use a different dropdown menu for each ingredient, in any order you like.**

**b) write the recipe that one should follow in order to cook what you see in the image. You should use the ingredients that you selected. Use a line break after each instruction you write.**



Choose the ingredients that are needed to cook what you see in the picture

empty	empty	empty
empty	empty	empty
empty	empty	empty
empty	empty	empty
empty	empty	empty
empty	empty	empty
empty	empty	empty

Write a recipe for the image above.



Figure 4: **User Study 2.** Recipe quality assessment form.

## QUALITATIVE EVALUATION FOR FOOD UNDERSTANDING METHODS

Predicting ingredients and writing recipes directly from food images are challenging tasks, which require strong visual recognition and reasoning models to understand food, its components and preparation instructions.

We have designed several methods to automatically solve these tasks, which we need to evaluate. In order to do so, we need your cooperation (and we are very grateful for it!).

In this test, **15 different images will be prompted, followed with a cooking recipe. For each image-recipe pair, you are required to decide whether the recipe is a plausible match to the image.**

Completing this form should take about 15 minutes. Note that partially completed forms cannot be submitted (it is mandatory to provide an answer for every image).

Don't forget to hit the **Submit** button at the end of the form when you are done.



**Title:** Mediterranean chicken pasta

**Ingredients:**

- Cheese
- Olive
- Pasta
- Tomato
- Pepper
- Chicken
- Oil
- Basil
- Clove

**Instructions:**

- Cook pasta as directed on package, omitting salt.
- Meanwhile, heat oil in large nonstick skillet on medium-high heat.
- Add chicken; cook and stir 5 min.
- Or until done.
- Add garlic; cook and stir 1 min.
- Stir in tomatoes, olives and basil; cook and stir 2 min.
- Or until heated through.
- Drain pasta; place in large bowl.
- Add chicken mixture; mix lightly.
- Sprinkle with cheese.

Is the recipe plausible for the image?



Figure 5: **Written Recipes.** Real, generated and human written recipes collected with our user study.




<div>Real</div> <div>Generated</div> <div>Written (1)</div> <div>Written (2)</div>			
	<b>Ingredients</b> hamburger, beans, tomato, soup, onion, macaroni, chili, sugar, ketchup, broth, butter, pepper <b>Instructions</b> -Combine all ingredients and cook over medium heat until potatoes are just tender. -Turn down heat to low and simmer at least 1.5 hours.	<b>Ingredients</b> pineapple, greens, stick, tomato, vinaigrette <b>Instructions</b> -Preheat greased grill to medium-high heat. -Grill fruit 3 min. -On each side or until lightly browned on both sides. -Cut fruit into 2-inch sticks; place in large salad bowl. -Add greens, jicama and tomatoes; toss lightly. -Drizzle with dressing just before serving.	<b>Ingredients</b> beef, onion, garlic, salt, pepper, parsley, basil, sugar, rice, broth <b>Instructions</b> -Brown the beef with spices and onion when browned add paste -Bring broth to boil add tomato sauce and spices let boil 3 minutes add rice bring back to boil cover and let sit off heat 7 minutes -Mix rice and sauce enjoy
	<b>Ingredients</b> chuck, tomato, water, soup <b>Instructions</b> -Brown ground chuck in a large dutch oven. -Drain any grease and add the crushed tomatoes, onion soup mix and water. -Simmer 5 minutes. -Add velveeta shells and cheese, mix well and serve with hot rolls.	<b>Ingredients</b> greens, cheese, tomato, ranch_dressing, chicken, italian_dressing <b>Instructions</b> -Toss greens with chicken, tomatoes, cheese and dressing in large bowl. -Add dressing; mix lightly.	<b>Ingredients</b> rice, onion, pepper, oil, salt, cheese, tomato, clove, broth <b>Instructions</b> -In a large skillet, heat oil over medium heat. -Add onion and garlic and cook until onion is translucent. -Add rice and cook until rice is lightly browned. -Add chicken broth, tomatoes, salt, pepper and cayenne pepper. -Bring to a boil. -Reduce heat and cover. -Simmer for 20 minutes. -Remove from heat and let stand covered for 5 minutes. -Sprinkle with cheese and serve.
	<b>Ingredients</b> tomato, chili, salt, beef, cheese, oil, pepper <b>Instructions</b> -Wash the tomato -Cut the tomato. -Put some oil in a pan -Add the beef -Add the tomato -Add the chili and the cheese -Add some salt and pepper	<b>Ingredients</b> tomato, spinach, sprouts, arugula, vinaigrette, zucchini <b>Instructions</b> -Chop the tomatoes. -Mix all the ingredients together. -Add the vinaigrette on top.	<b>Ingredients</b> rice, meat, parsley, salt, oil, cheese <b>Instructions</b> -Cook the minced meat in a big pan. Put salt and oil into it -When it is cooked, add the rice and water so it can be boiled -Boil it for 20 minutes - When all the water has evaporated, add some cheese and parsley to enhance the flavor.
	<b>Ingredients</b> beans, seasoning, tomato, beef, spices, hot_sauce, cheese, onion <b>Instructions</b> -Cook the beef in a frying pan till brown. -Dice the onions. -Drain the beans. -Take a slow-cooker and put the beef, tomatoes, beans and onions. -Add the seasoning and spices. -Turn the slow-cooker on medium and cook for around 6 hrs.	<b>Ingredients</b> arugula, tomato, kale, zucchini, onion, oil <b>Instructions</b> -Wash and rinse the kale -Wash, cut and slice the onion, the yellow zucchini and the cherry tomatoes -Add all the ingredients in a bowl and mix -Add some olive oil and your favorite vinagrette	<b>Ingredients</b> rice, parsley, broth, meat, salt, water <b>Instructions</b> -Put the sauteed meat in a casserole with hot oil. -After it is quite cooked, add the rice. -Add the water. -Add the broth. -Add salt -Let it cook until the rice it is done. -Add the parsley

Figure 6: **Dine Out Study.** Generated recipes for food images taken by authors.



**Title:** Chicken enchiladas

**Ingredients:**

cheese, cream, tortilla, onion, enchilada\_sauce, chili, chicken, pepper, salt, milk, tomato, cumin

**Instructions:**

- Preheat oven to 350 degrees.
- In a large skillet, saute onion and green pepper in the oil until tender.
- Add chicken, tomatoes, enchilada sauce, green chilies, cumin, salt, pepper, and chili powder.
- Simmer for 10 minutes.
- In a large bowl, combine sour cream, milk, and cheese.
- Stir in chicken mixture.
- Spoon about 1/3 cup of chicken mixture down the center of each tortilla.
- Roll up and place seam side down in a greased 9x13 baking dish.
- Pour enchilada sauce over enchiladas.
- Sprinkle with cheese.
- Bake for 20 minutes.

(a)



**Title:** Potato soup with dill

**Ingredients:**

egg, pepper, potato, cream, parsley, butter, salt, milk, onion, dill, water

**Instructions:**

- In a large pot, bring the potatoes, water, salt, and pepper to a boil.
- Reduce the heat to low and simmer until the potatoes are tender, about 20 minutes.
- Drain the potatoes and return them to the pot.
- Add the butter and milk and mash until smooth.

(b)



**Title:** Prosciutto and egg frittata

**Ingredients:**

egg, pepper, cheese, bacon, oil, salt, prosciutto

**Instructions:**

- Preheat oven to 350 degrees.
- In a large skillet, heat olive oil over medium heat.
- Add prosciutto and cook until crisp, about 5 minutes.
- Remove bacon from skillet and set aside.
- In a medium bowl, whisk eggs, salt and pepper.
- Add prosciutto and stir to combine.
- Pour egg mixture into skillet and cook until eggs are set, about 5 minutes.

(c)



**Title:** Mesclun salad with pomegranate and avocado

**Ingredients:**

oil, tomato, greens, mayonnaise, vinegar, mesclun, pepper, seeds, chili, juice

**Instructions:**

- In a small bowl, whisk the mayonnaise with the lime juice, vinegar and chile powder.
- Season with salt and pepper.
- In a large bowl, toss the mesclun with the tomatoes, avocado, tomatoes and sesame seeds.
- Drizzle with the dressing and serve.

(d)