



UNIVERSITAT DE
BARCELONA



MADiMa 2017



Food Ingredients Recognition through Multi-label Learning

Marc Bolaños, Aina Ferrà and Petia Radeva

Motivation

THE WORLD IS GETTING FATTER



HOW DO I KNOW WHETHER I AM OVERWEIGHT?

Calculate your body mass index (BMI) using this formula

$$\text{BMI} = \frac{\text{weight (kg)}}{\text{height}^2 \text{ (m}^2\text{)}}$$



OBESITY KILLS!

7 common diseases due to obesity:

- Arthritis
- Cancer
- Infertility
- Heart Diseases
- Back Pain
- Diabetes
- Stroke

OBESITY IS KILLING PREVENTABLE THE WORLD

A B C TO OBESITY PREVENTION

SIMPLE RULES TO STAY IN SHAPE

A dopt New Healthy Habits



B alance Your Calorie Intake



C ontrol Your Weight Gain



source: World Health Organization ©2014 Health Buzz www.healthbuzz.asia

Motivation



Motivation

INPUT



Computer
Vision
Algorithm



List of
Ingredients



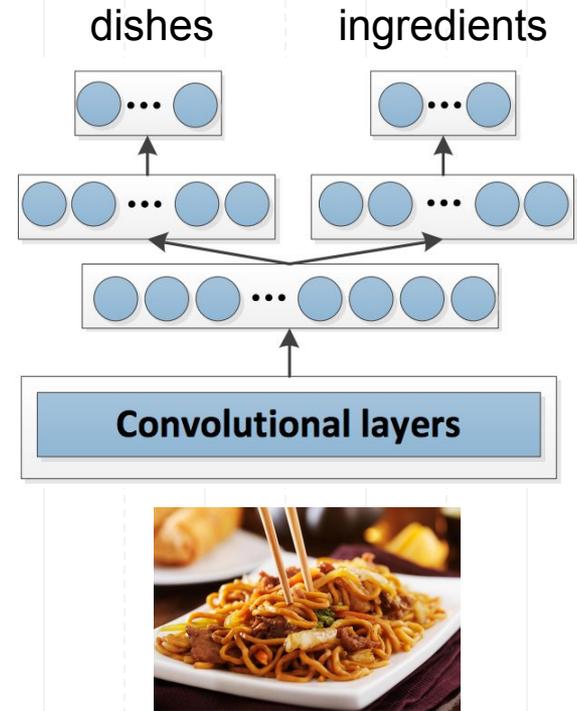
Nutritional
Composition

OUTPUT



Related Work

Only one work has been proposed for ingredients recognition.



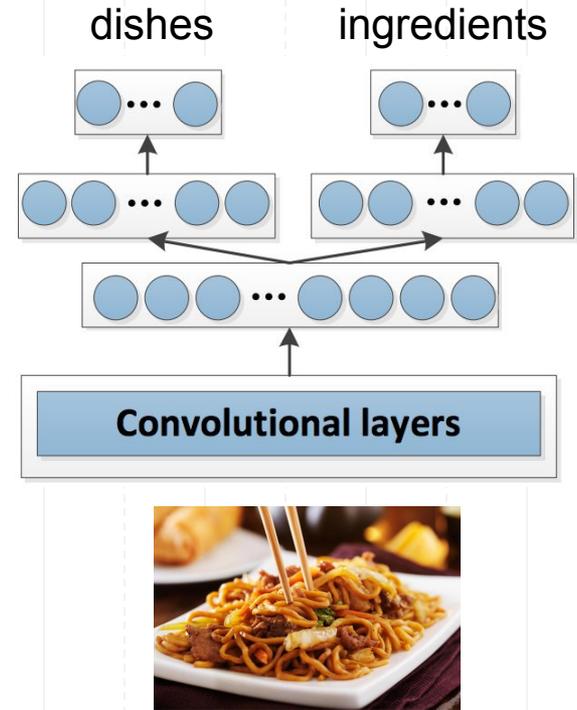
Jingjing Chen and Chong-Wah Ngo. "Deep-based ingredient recognition for cooking recipe retrieval". In Proceedings of the 2016 ACM on Multimedia Conference, pages 32–41. ACM, 2016.

Related Work

Only one work has been proposed for ingredients recognition.

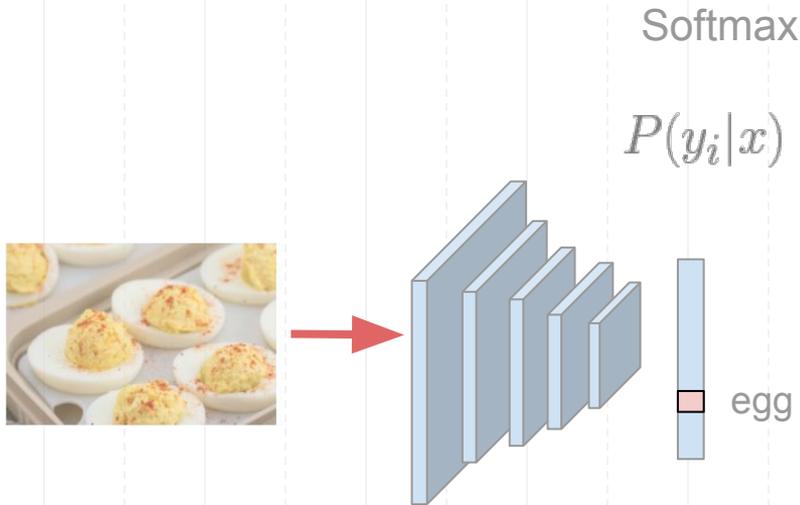
Handicaps of their proposal:

- Both dish and ingredients information is needed. Not applicable for never-seen recipes.
- Applicable to visible ingredients only.



Jingjing Chen and Chong-Wah Ngo. "Deep-based ingredient recognition for cooking recipe retrieval". In Proceedings of the 2016 ACM on Multimedia Conference, pages 32–41. ACM, 2016.

Conventional CNN for classification



Softmax

$$P(y_i|x) = \frac{\exp^{f(x)_i}}{\sum_i \exp^{f(x)_j}}$$

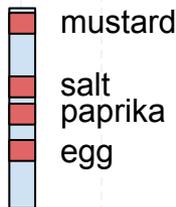
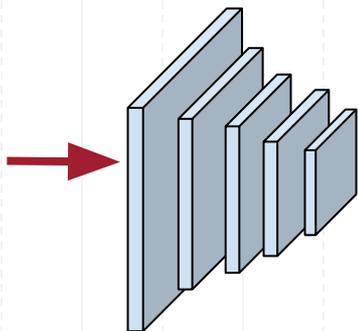
loss



Categorical cross-entropy

$$L_c = - \sum_x \log(P(\hat{y}_x|x))$$

Our proposal: Adaptation for Multi-label recognition



Sigmoid

$$P(y_i|x) = \frac{1}{1 + \exp^{-f(x)_i}}$$

loss

Binary cross-entropy

$$L_b = - \sum_x \sum_i^N (\hat{y}_x^i \cdot \log(P(y_i|x)) + (1 - \hat{y}_x^i) \cdot \log(1 - P(y_i|x)))$$

Datasets - Ingredients101

Dataset complementary to Food101:

- 101 classes / dishes
- 1000 images per class

A recipe for each class was downloaded from *Yummly*™ resulting in a list of ingredients per class and a total of 446 unique ingredients.

Carrot Cake



Ingredients: 'salt', 'butter', 'all-purpose flour', 'large eggs', 'vanilla extract', 'baking powder', 'carrots', 'granulated sugar', 'powdered sugar', 'baking soda', 'brown sugar', 'ground cinnamon', 'canola oil', 'cream cheese', 'sour cream', 'ground nutmeg', 'chopped pecans', 'unsweetened applesauce',

Baby Back Ribs



Ingredients: 'barbecue sauce', 'baby back ribs', 'chips', 'barbecue rub',

Datasets - Ingredients101

Dataset complementary to Food101:

- 101 classes / dishes
- 1000 images per class

A recipe for each class was downloaded from *Yummly* resulting in a list of ingredients per class and a total of 446 unique ingredients.

Carrot Cake



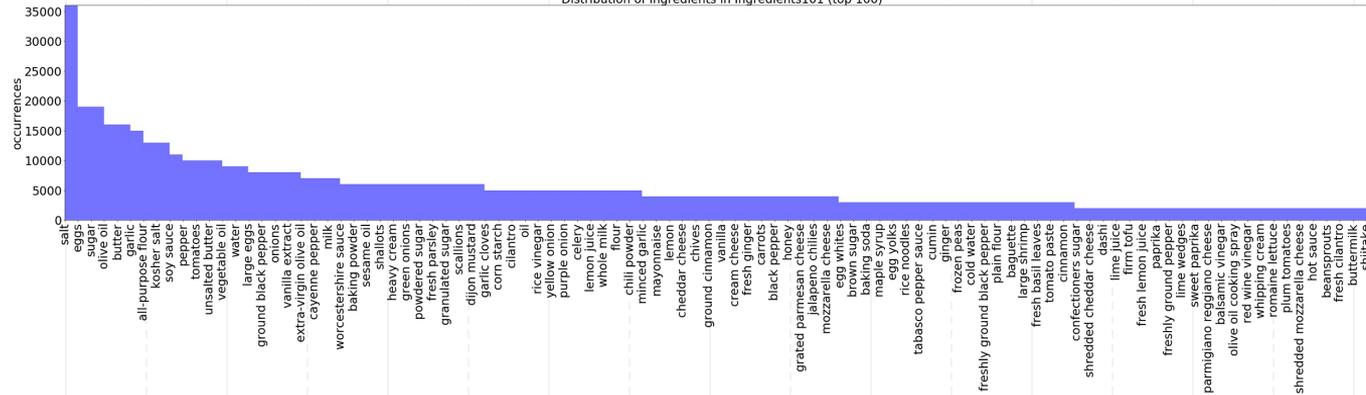
Ingredients: 'salt', 'butter', 'all-purpose flour', 'large eggs', 'vanilla extract', 'baking powder', 'carrots', 'granulated sugar', 'powdered sugar', 'baking soda', 'brown sugar', 'ground cinnamon', 'canola oil', 'cream cheese', 'sour cream', 'ground nutmeg', 'chopped pecans', 'unsweetened applesauce',

Baby Back Ribs



Ingredients: 'barbecue sauce', 'baby back ribs', 'chips', 'barbecue rub',

Distribution of Ingredients in Ingredients101 (top 100)



Datasets - Recipes5k

Ingredients Simplification

Two problems arise when dealing with too many labels:

- The amount of training samples needed for learning them.
- The ambiguity and minor differences between them.

We propose a simplified version by applying a simple removal of overly-descriptive particles (e.g. 'sliced' or 'sauce'). **Simplifying the 3,213 ingredients into 1,013.**



large eggs → egg
lemon zest → lemon
meyer lemon juice → lemon
unsalted butter → butter
boiling water → water

Results - Ingredients101

	Validation			Test		
	Prec	Rec	F_1	Prec	Rec	F_1
Random prediction	2.05	2.01	2.03	2.06	2.01	2.04
InceptionV3 + Ingredients101	80.86	72.12	76.24	83.51	76.87	80.06
ResNet50 + Ingredients101	84.80	67.62	75.24	88.11	73.45	80.11

Results - Ingredients101



Dish: lasagna

Prediction: 'eggs', 'garlic', 'grated parmesan cheese', 'black pepper', 'fennel seeds', 'shredded mozzarella cheese', 'diced tomatoes', 'chopped onion', 'lasagna noodles', 'pork sausages', 'ricotta cheese', 'italian seasoning', 'tomato sauce',

GT: 'eggs', 'garlic', 'grated parmesan cheese', 'black pepper', 'fennel seeds', 'shredded mozzarella cheese', 'diced tomatoes', 'chopped onion', 'lasagna noodles', 'pork sausages', 'ricotta cheese', 'italian seasoning', 'tomato sauce',



Dish: seaweed_salad

Prediction: 'tomatoes', 'rice vinegar', 'seaweed', 'sesame seeds',

GT: 'tomatoes', 'rice vinegar', 'seaweed', 'sesame seeds',



Dish: fried_calamari

Prediction: 'kosher salt', 'all-purpose flour', 'vegetable oil', 'lemon', 'mayonnaise', 'parsley leaves', 'masa harina', 'calamari', 'parsley', 'meyer lemons',

GT: 'kosher salt', 'all-purpose flour', 'vegetable oil', 'lemon', 'mayonnaise', 'parsley leaves', 'masa harina', 'calamari', 'parsley', 'meyer lemons',

Results - Ingredients101



Dish: prime_rib

Prediction: 'olive oil', 'kosher salt', 'minced garlic', 'thyme', 'peppercorns', 'rosemary', 'rib-eye roast',

GT: 'olive oil', 'kosher salt', 'minced garlic', 'thyme', 'peppercorns', 'rosemary', 'rib-eye roast',



Dish: caesar_salad

Prediction: 'salt', 'extra-virgin olive oil', 'dijon mustard', 'freshly ground black pepper', 'red wine vinegar', 'dried mixed herbs', 'toasted pine nuts', 'beets', 'gorgonzola', 'baby spinach',

GT: 'salt', 'garlic', 'pepper', 'dijon mustard', 'worcestershire sauce', 'lemon juice', 'romaine lettuce', 'croutons', 'plain greek yogurt', 'parmesan cheese', 'anchovy paste',



Dish: chicken_curry

Prediction: 'salt', 'sugar', 'vegetable oil', 'ground black pepper', 'yellow onion', 'corn starch', 'garlic cloves', 'fresh ginger', 'frozen peas', 'chopped fresh cilantro', 'boneless skinless chicken breasts', 'low sodium chicken broth', 'greek yogurt', 'curry powder',

GT: 'salt', 'sugar', 'vegetable oil', 'ground black pepper', 'yellow onion', 'corn starch', 'garlic cloves', 'fresh ginger', 'frozen peas', 'chopped fresh cilantro', 'boneless skinless chicken breasts', 'low sodium chicken broth', 'greek yogurt', 'curry powder',

Results - Recipes5k

	Validation			Test		
	Prec	Rec	F_1	Prec	Rec	F_1
→ Random prediction	0.33	0.32	0.33	0.54	0.53	0.53
InceptionV3 + Ingredients101				23.80	18.24	20.66
ResNet50 + Ingredients101				26.28	16.85	20.54
→ InceptionV3 + Recipes5k	36.18	20.69	26.32	35.47	21.00	26.38
ResNet50 + Recipes5k	38.41	19.67	26.02	38.93	19.57	26.05
→ Random prediction	6.27	6.29	6.28	6.14	6.24	6.19
InceptionV3 + Ingredients101				44.01	34.04	38.39
ResNet50 + Ingredients101				47.53	30.91	37.46
→ InceptionV3 + Recipes5k	56.77	31.40	40.44	55.37	31.52	40.18
ResNet50 + Recipes5k	56.73	28.07	37.56	58.55	28.49	38.33
→ InceptionV3 + Recipes5k simplified	53.91	42.13	47.30	53.43	42.77	47.51

SIMPLIFIED
INGREDIENTS

Results - Recipes5k



Dish: churros

Prediction: 'salt', 'water', 'granulated sugar', 'oil', 'cinnamon', 'ground cinnamon',

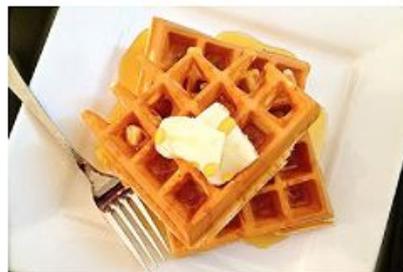
GT: 'sugar', 'water', 'unsalted butter', 'kosher salt', 'large eggs', 'vanilla extract', 'flour', 'cinnamon', 'ground cinnamon', 'canola oil',



Dish: churros

Prediction: 'salt', 'water', 'sugar', 'oil', 'cinnamon', 'cinnamon',

GT: 'sugar', 'water', 'butter', 'salt', 'egg', 'vanilla', 'flour', 'cinnamon', 'cinnamon', 'oil',



Dish: waffles

Prediction: 'salt', 'eggs', 'sugar', 'all-purpose flour', 'milk', 'large eggs', 'baking powder', 'granulated sugar', 'melted butter', 'blueberries',

GT: 'salt', 'sugar', 'all-purpose flour', 'milk', 'baking powder', 'vanilla', 'cooking oil', 'beaten eggs',



Dish: waffles

Prediction: 'salt', 'egg', 'sugar', 'flour', 'milk', 'egg', 'baking', 'sugar', 'butter', 'blueberries',

GT: 'salt', 'sugar', 'flour', 'milk', 'baking', 'vanilla', 'oil', 'egg',

Results - Recipes5k



Dish: eggs benedict

Prediction: 'salt', 'eggs', 'butter', 'unsalted butter', 'lemon juice', 'egg yolks', 'fresh lemon juice', 'parsley', 'white vinegar', 'large egg yolks', 'english muffins', 'canadian bacon',

GT: 'salt', 'unsalted butter', 'large eggs', 'large egg yolks', 'lemon zest', 'english muffins', 'ham', 'boiling water', 'fresh marjoram', 'meyer lemon juice',



Dish: eggs benedict

Prediction: 'salt', 'egg', 'butter', 'butter', 'lemon', 'egg', 'lemon', 'parsley', 'vinegar', 'egg', 'english muffin', 'bacon',

GT: 'salt', 'butter', 'egg', 'egg', 'lemon', 'english muffin', 'ham', 'water', 'marjoram', 'lemon',



Dish: french onion soup

Prediction: 'sugar', 'onions', 'all-purpose flour', 'black pepper', 'yellow onion', 'balsamic vinegar', 'dry white wine', 'baguette', 'french bread', 'bay leaves', 'beef broth', 'gruyere cheese', 'fresh thyme', 'red wine', 'sweet onion', 'beef stock',

GT: 'salt', 'sugar', 'pepper', 'unsalted butter', 'flour', 'yellow onion', 'french bread', 'gruyere cheese', 'fresh thyme', 'beef stock', 'sherry wine',



Dish: french onion soup

Prediction: 'sugar', 'onion', 'flour', 'black pepper', 'onion', 'balsamic vinegar', 'white wine', 'baguette', 'bread', 'bay', 'beef', 'cheese', 'thyme', 'red wine', 'onion', 'beef',

GT: 'salt', 'sugar', 'pepper', 'butter', 'flour', 'onion', 'bread', 'cheese', 'thyme', 'beef', 'sherry',

Neurons' Activations

Ingredient activation: lettuce



Dish: pad thai

Dish: hamburger

Dish: hamburger

Dish: tiramisu

Dish: tiramisu

Dish: hamburger

Dish: hamburger

Dish: hamburger

Dish: hamburger

Dish: tiramisu

Ingredient activation: ketchup



Dish: hamburger

Dish: hamburger

Dish: hamburger

Dish: beef tartare

Dish: hamburger

Dish: pulled pork sandwich

Dish: pulled pork sandwich

Dish: cannoli

Dish: macarons

Dish: hamburger

Ingredient activation: all-purpose flour



Dish: donuts

Dish: churros

Dish: churros

Dish: filet mignon

Dish: chicken curry

Dish: churros

Dish: donuts

Dish: breakfast burrito

Dish: churros

Dish: churros

Neurons' Activations

Ingredient activation: butter



Ingredient activation: granulated sugar



Ingredient activation: mayonnaise



Neurons' Activations

Ingredient activation: soy sauce



Dish: pad thai

Dish: takoyaki

Dish: takoyaki

Dish: takoyaki

Dish: scallops

Dish: gyoza

Dish: tuna
tartare

Dish: takoyaki

Dish: gyoza

Dish: takoyaki

Ingredient activation: garlic



Dish: pad thai

Dish: gyoza

Dish: gyoza

Dish: pulled
pork
sandwich

Dish: pulled
pork sandwich

Dish:
hamburger

Dish: pulled
pork sandwich

Dish: gyoza

Dish: gyoza

Dish: bread
pudding

Conclusions

We have proposed:

- Model suitable for ingredients recognition through multi-label learning.
- Two datasets for ingredients recognition benchmarking.

Advantages of our proposal with respect to the state of the art:

- Straightforward model applicable to any highly performing CNN.
- Dish/class information is not used for learning. Implying that the ingredients can be inferred from never-seen dishes.
- Can directly learn ingredients' representation from visual appearance.
- Can predict invisible ingredients implicitly.

THANK YOU FOR YOUR ATTENTION



www.ub.edu/cvub/marcbolanos



marc.bolanos@ub.edu