Food9K: Social Media Food Image Detection and Localization

Professor Mohammad Akbari
Zahra Golpayegani

Introduction

- What is food consumption behavior analysis?

- What is the importance of food consumption behavior analysis?

Problem Statement

- How to create an in-the-wild food dataset that can represent food consumption behaviors?
- How can we detect and localize food items within an image?

Contributions

- Food9K - a dataset of 35 food classes crawled from social media with ~9000 food images and their metadata
- Food detection and localization YOLO model trained on Food9K

Literature Review - Food Datasets

DATSET NAME	#CLASSES	#IMAGES	SOURCE
ETHZ Food-101	101	101000	foodspotting.com
UPMC Food-101	101	100000	Google Images Search
UNICT-FD889	889	3583	smart phone
FooDD	23	3000	camera
Menu-Match	41	646	social media
UNIMIB2015	15	2000	smart phone
UNIMIB2016	73	1027	smart phone
Instagram800K	43	808964	Instagram
Food524DB	524	247636	existing datasets
ChineseFoodNet	208	192000	web
UECFOOD-256	256	341517	web

Literature Review - Food Dictionaries

- What do we mean by “Food Dictionary”?
- Main approaches:

  > Manually analyze restaurant menus
  > Use a food vocabulary list
  > Asking experts

Literature Review - Visual Food Recognition and Localization

- Single-label food recognition vs multi-label food recognition
- CNN-based approaches outperformed conventional methods
- Several CNN architectures has been studied, e.g., Inception for food recognition
- A heat map of probabilities for generating bounding boxes proposals

Food9K Dataset

- Food dictionary: Created by manually analyzing DOHMH MenuStat dataset with 335 food items.

- Data collection: Images that contained at least one keyword from the food dictionary were downloaded from social media with their metadata.

- Data cleaning: 1) Food/non-food filtering 2) food classes with less than 2000 images were removed from the dataset

- Data labeling: Run a pre-trained YOLO model on Food9K to have bounding box suggestions and verified them manually

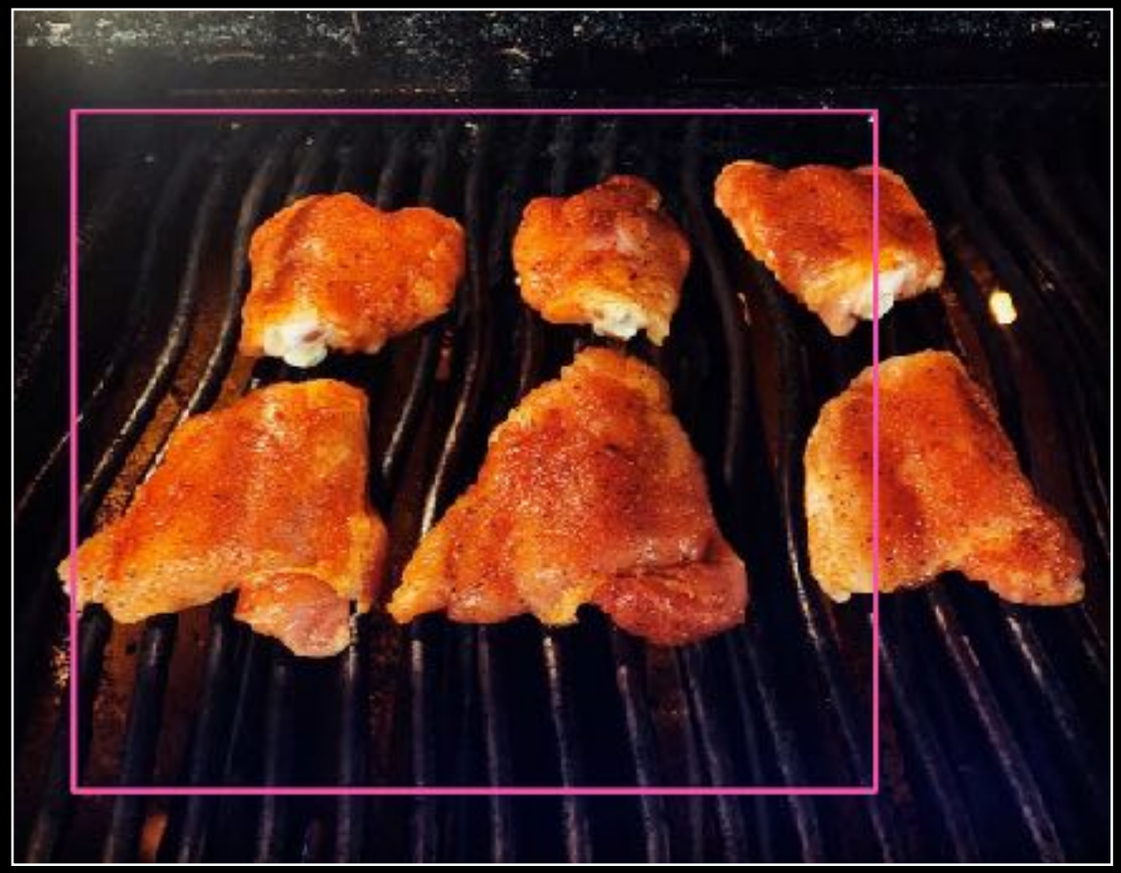
Food9K Dataset - The Labeling Process

- Train a YOLO model on the UECFOOD-256 dataset
- Use the model to help with labeling Food9K

Model Name	Precision	Recall	F1-score	GIoU
UECFOOD-256	0.517	0.55	0.516	0.924

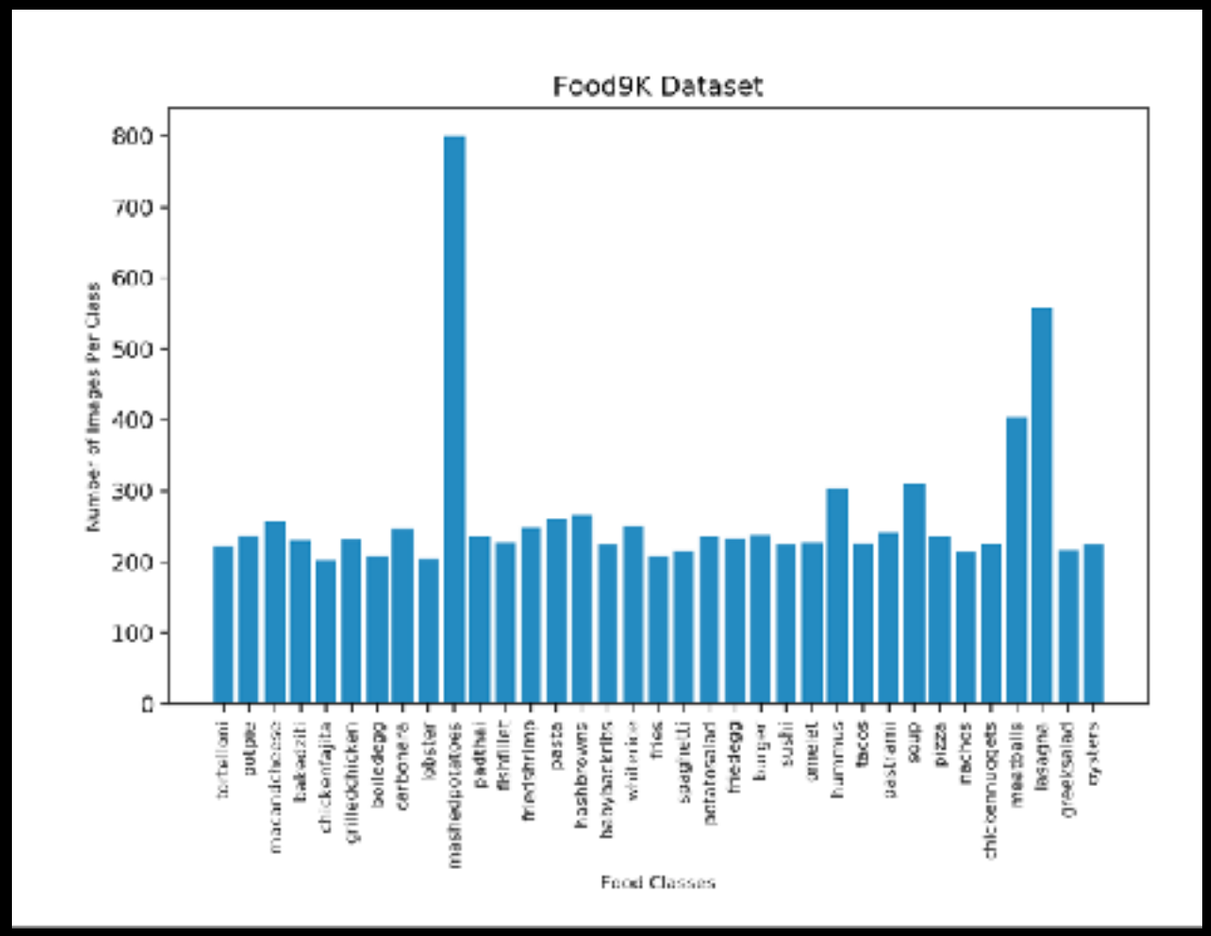
YOLO model performance on UECFOOD-256 dataset

Bounding Box Proposals



Examples of UECFOOD-256 model performance on Food9K dataset

Food9K at a Glance

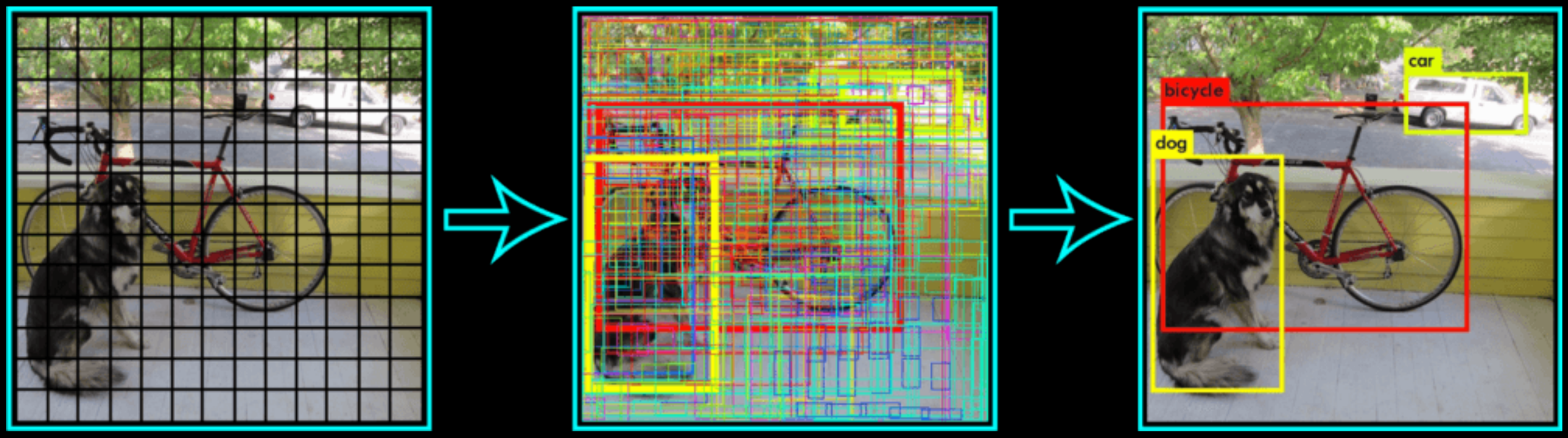



```
13,"Verdict: Good. Not great. If you watched our LIVE you saw us making this instant pot
baked ziti. Omg. So excited for it and so happy with how easy it was. Flavor wise, the
recipe is very lacking in spices and flavor. If you know anything about us...we are NOT
salt fans but boy does this need some salt!  I sprinkled a little extra in my second
helping and it helped. All in all a good experiment. Now that we are getting better with
the IP we can feel more comfortable making our own adjustment. •
•
•
•
•
#instantpot #instantpotrecipes #bakedziti #bakedzitipasta #pasta #pastafordinner
#garlicbread #foodporn #foodphotography #foodpics #instantpotmeals
#instantpotindiancooking #welovefood #nom #foodie #yum #cookingwiththediefs",2020-02-29
00:34:41,2,B9IX-PvHsxp,730
```

Metadata sample

Training the Model

- What is YOLO?



Results

Model Name	Precision	Recall	F1-score	GIoU
Food9K	0.591	0.692	0.633	1.67

YOLO model performance on Food9K dataset

Future Works

- Extend Food9K dataset
- Improve model performance
- Draw insights about food consumption behavior from social media data (Food9K dataset)
- Design a diet assessment application using the trained model

References

- Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

- Bossard, Lukas, Matthieu Guillaumin, and Luc Van Gool. "Food-101–mining discriminative components with random forests." European conference on computer vision. Springer, Cham, 2014.

- Francesco Ragusa, Valeria Tomaselli, Antonino Furnari, Sebastiano Battiato, and Giovanni M. Farinella. 2016. Food vs Non-Food Classification. In Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management (MADiMa '16). Association for Computing Machinery, New York, NY, USA, 77–81. https://doi.org/10.1145/2986035.2986041

- Jalal, Mona, et al. "Scraping social media photos posted in Kenya and elsewhere to detect and analyze food types." Proceedings of the 5th International Workshop on Multimedia Assisted Dietary Management. 2019.

- Continued (See the paper for a complete list of references)