The George Washington University

Columbian College of Arts & Sciences

DATS 6303 – Deep Learning

Professor AmirJafari,PhD

Team 3: Chirag Lakhanpal, Tejas Rawal & Yashwant Bhaidkar

AI-Driven Food Item Recognition for Enhanced Dietary Tracking

Project Proposal

Problem Selection and Rationale

The project aims to create a model that can identify and categorize individual food items from images. This is a critical task for personal health tracking and medical dietary studies. This task was chosen because of its potential to improve the accuracy of food intake documentation, which is essential for nutritional research and personal health management. We then aim to calculate the calories in each food item in order to track calorie intake per meal.

Database/Dataset

The selected dataset for this project is the AIcrowd Food Recognition Challenge dataset. The data comprises 24,120 training images with 39,328 annotations and 1,269 validation images with 2,053 annotations, formatted in the MS-COCO style. This volume of data, combined with ongoing dataset expansion, is sufficiently large for training an effective deep-learning model.

Deep Learning Model

The project will utilize a Convolutional Neural Network (CNN), specifically the Mask R-CNN architecture, due to its success in object detection and instance segmentation tasks. Given the nature of the food recognition problem, customizations may be necessary for handling the dataset's unique characteristics, such as varying textures and food groupings. Post using standard models we will employ pre-trained models like YOLO, EfficientNet, Xception, and RetinaNet to check for better accuracy.

Implementation Framework

The PyTorch framework will be the main tool for implementing the neural network. PyTorch provides a dynamic computation graph that is useful for variable-length inputs, which are common in image-based datasets. Additionally, its extensive library of tools and active community support make it an ideal choice for this project.

Reference Materials:

Our project will be based on a variety of learning resources to ensure a thorough understanding of image recognition and segmentation. We will review a variety of technical documentation, online courses, and interactive forums that provide insights into the deployment of Convolutional Neural Networks for complex image-based tasks. Additionally, we intend to mine the collective knowledge found in online tech communities like Stack Overflow and GitHub, where real-world problem-solving examples are plentiful. This rich tapestry of learning materials will serve as the intellectual foundation for our project's development.

Performance Metrics

Model performance will be evaluated using the Intersection over Union (IoU) metric, as suggested by the challenge guidelines. Precision and recall at an IoU threshold of 0.5 will be the primary focus, with Average Precision (AP) and Average Recall (AR) serving as aggregate performance indicators.

Project Schedule (Tentative)

- Week 1: Literature review and dataset exploration model selection and baseline implementation.
- Week 2 Model training and validation.
- Week 3 Performance tuning and model optimization and evaluation on the test set and documentation.
- Week 4 Finalizing report and preparing a presentation of our findings.