# Computer Vision 2026 Project [SC & AI]

# Food/Fruit Recognition and Calorie Estimation

Food/Fruit Recognition and Calorie Estimation is a crucial task to automate dietary tracking, enhance nutritional monitoring, and simplify personalized health management. For instance, a user or health specialist could leverage this technology to automatically identify different food and fruit types within a meal photograph and accurately estimate the calorie count. This information can then be used to provide real-time feedback on daily intake, inform dietary choices for weight management or specific health conditions, and monitor consumption patterns over time.

Project Objectives:

1. Apply Image preprocessing or Feature Extraction techniques where needed.
2. There are two identification stages required in the project: the first stage is to determine whether the given image is food or fruit, the second stage determines the category of food or fruit.
3. Train a classification model to determine whether the given image is food or fruit (This is the first stage).
4. For the second stage (Fruit), train a classification model to determine the fruit category.
5. For the second stage (Food), train a one/few shot learning model (e.g., Siamese network) to recognize the food category. This approach is necessary because, unlike static Fruit categories, Food categories frequently encounter novel dishes and ingredients, requiring the model to adapt quickly to unseen classes at test time.
6. Train a binary segmentation model, that segments the fruit regardless of its type from the background.

7. Train a second multi-stage segmentation model that segments each pixel to one of 31 classes: background and the 30-fruit class.

   **Minimum Requirements:**
   a. You must train at least one classification model for the first stage.
   b. You must train at least one classification model to determine the fruit type and a one/few shot learning model (e.g., Siamese) for recognizing the food type for the second stage.
   c. You must train at least one binary segmentation model and at least one multi-class segmentation model.

# Dataset Description

## The dataset for the project can be found [here].

The dataset consists of Food and Fruit directories, The Food directory consists of Train and Validation folders with the corresponding calorie per gram for each category, and the Fruit directory consists of Train and Validation folders, where each fruit category/folder within training and validation has image folder and mask folder.

In Part A (1st stage classification), you are only required to train a model responsible for classifying whether the given image is food or fruit regardless of the type, so all the images within the categories belonging to the food directory should have one label (**Food**), while all the images within the categories belonging to fruit directory should have one label (**Fruit**) You should also report the validation accuracy on all the validation images belonging to Food and Fruit categories.

In Part B – 2nd stage Food recognition (**one/few shot recognition**), you are required to train a one/few shot learning model, which is responsible for determining the food category type.

In Part C – 2nd stage Fruit recognition, you are required to train a model, responsible for classifying the Fruit type.

In Part D – Fruits Binary Semantic Segmentation, you will train a model to perform binary semantic segmentation of various fruits. The objective is to predict a mask where the fruit (foreground) is represented by white pixels, and the background is represented by black pixels. The ground truth data structure includes an Images directory and a corresponding Mask directory within each fruit's category.

In Part E – Fruits Multi-Class Semantic Segmentation, you will train a model to perform multi-class segmentation of fruits. The segmentation must distinguish between 31 total classes: 30 unique fruit types and the background. The ground truth mask for each image must accurately differentiate and label every pixel based on its corresponding category (one of the 30 fruit types or the background).

## Practical Exam Project Deliverables:

1. Food or Fruit Classification part (Part A - 1st stage) (Deliver Code).
2. Food Recognition/Verification (Part B - 2nd stage) (Deliver Code).
3. Fruit Classification (Part C – 2nd stage) (Deliver Code).
4. Fruit Binary Segmentation (Part D) (Deliver Code).
5. Fruit Multi-Class Segmentation (Part E) (Deliver Code).
6. If you trained the deep learning/Siamese models/sementic segmentation models using a notebook, you must deliver the notebook with the output cell saved displaying the training logs. If you trained the model using IDE (i.e Pycharm). **You must deliver screenshots of the training process**.
7. **Individual Test Scripts**: Create test scripts for each part; The first should take an image and classify it to either food or fruit (**For Part A**), The test script in **Part B** should handle two cases, the first case is to recognize the food type by comparing each given test image to the training categories, while the second case, we will give you some images to new food categories unseen in the validation and training data (N images) and will give you a reference image (Anchor), and

you should compare this Anchor image with all other given unseen images (N images) and recognize to whom food category does it belong to (most similar to), or if it isn't similar to any of the images then you can say that there is no match. The test script in **Part C** should classify the given image into one of the fruit classes. The test script in **Part D** should generate a pixel-wise mask for each given image where white corresponds to fruit and black for background. The test script in **Part E** should generate a pixel-wise mask for each given image where each fruit should have a different color depending on its type and black for background. You need to output a text file containing the predicted output for each task (Three text files for each of the parts: A,B,C) and save the predicted segmentation masks for Parts D,E. These saved files will be in addition to printing the output and displaying it in the code.

8. **Integrated Test Pipeline**: You will be given a set of images, you need to classify whether each is food or fruit, if food you need to determine its type and total number of calories, if fruit you need to determine its type and total number of calories, as well as you need to save the binary segmentation result and the multi-class segmentation result in case of fruit. You need to output a text file with the first line specifying whether the given image is food or fruit, and the second line will be the recognized type of food or fruit, the third line will be the total number of calories (**you will be given the image in test with name imgx_120g for example**) and you need to extract the total number of grams (e.g., 120) to multiply them by the calories within 1g of the recognized type to get the total calories), as well as in case of fruit, you need to save two images corresponding to binary segmentation result and multi-class segmentation result as mentioned above.
9. A Report that includes description of:
    - Your data preparation process.
    - Brief description of the models and techniques used in each task.

- Training and Testing times for each model.
- Image Classification training and validation accuracy.
- Provide screenshots of the validation sets classification with visualization.
- Image recognition (one/few shot learning model) training and validation accuracies.
- Image segmentation training and validation accuracies along with screenshots of segmented samples for binary and multi-class segmentation scenarios.