

人工智能应用实践 课程结题报告

项目题目：	<b>Restauraut+</b> 食物识别分析与营养规划系统
小组成员：	姓名 倪浚桐 学号 202022161224
	姓名 欧阳坤 学号 202000161069
	姓名 曾爽 学号 202018161187

**项目背景描述：**

项目以落地性强、准确度高为主要宗旨。

对于餐厅而言。目前，由于使用收银机，餐厅的付款流程仍然是人工的且效率低下的。收银员会检查顾客点了什么食物，然后在收银台上进行结算。效率并不高。因此，食物识别设备和自动食物价格估算可以解决这些问题。食物识别分析旨在优化餐厅付款流程，并使用计算机视觉方法自动估算食物价格。传统的方法有餐盘识别价格计算法，该方法通过设置价格区间，不同的价格对应不同颜色的餐盘，因此算法只需要有能力识别各种颜色的餐盘即可，这样做优点是准确率高、算法设计简单，缺点是价格只能设定为一些固定的数值，这加大了食堂工作人员定价的难度，缺乏灵活性。而基于深度学习的，直接对食物进行识别检测，则只需对于每一种食物进行定价即可，这样的方法灵活性好，应用范围更广。

对于消费者而言。我们知道，食物含有不同的卡路里。在一顿饭中，消费者通常不知道所供应的食物中有多少卡路里。消费者需要通过食物识别分析，在一张食物照片中检测出提供的多种食物并根据数据库卡路里数据计算这一顿饭的总卡路里数。同时，消费者更希望合理的规划每天摄入的卡路里总量，因此我们还提供了营养规划系统。现在也有一部分智能手机软件能够实现卡路里计算的功能，但所有这些软件都需要用户手动上传食物重量，然后软件才能计算卡路里。这样的方式大幅度减低用户记录效率，有可能使得用户记录兴趣下降，导致记录习惯难以坚持，达不到软件预期效果。而基于深度学习的卡路里预测则无需用户手动上传重量，用户仅需要拍摄一张图片，算法则会根据图片预测出其质量以及含有的卡路里等营养成分含量。

因此，该项目能够改进过去的传统方法的一些弊端，实现更加准确、更加易用、更人性化、更具落地应用价值的系统功能。

项目的名称为“Restauraut+ 食物识别分析与营养规划系统”，该系统包含细粒度食物识别模块、食物价格匹配模块、食物卡路里分析模块以及营养规划模块。该项目以落地性强、准确度高为主要宗旨，服务对象包括餐厅、消费者等，可以广泛应用于餐厅收银、个人营养规划等实际用途。

#### 项目技术要点与运行环境：

##### 运行环境：

Python 3.7.13

Pytorch 1.7.1

##### 技术要点：

使用 **YOLOv5<sup>1</sup>** 进行食物识别；根据外部知识<sup>2</sup>对识别的食物得出价格；使用 **ResNet<sup>3</sup>** 等深度卷积神经网络在 **Nutrition5k** 数据集<sup>45</sup> 上进行训练以预测卡路里等营养成分含量；营养规划则将模型预测结果进行分类，使用 **ChatGPT<sup>6</sup>** 生成相应的评价及饮食搭配建议，由于直接将预测的营养成分输入到 ChatGPT 并不能很好地引导其给出评价，故我们会人工处理输入的文本内容。

难点在于：该项目大多是企业级别的开发项目，完全没有相关的开源项目实现，所以该项目完全由我们自己动手完成。

#### 项目目标与完成情况：

##### 项目目标：

最终的项目目标是针对给定一段视频输入，待视频中餐盘稳定后，根据图像进行多个模块的处理，输出食物选框及一段食物分析文字。食物分析文字会给出食物的价格信息和卡路里的含量“高中低”等

<sup>1</sup> <https://github.com/ultralytics/yolov5>

<sup>2</sup> <https://www.numbeo.com/food-prices/>

<sup>3</sup> <https://arxiv.org/abs/1512.03385>

<sup>4</sup> <https://arxiv.org/abs/2103.03375>

<sup>5</sup> <https://github.com/google-research-datasets/Nutrition5k>

<sup>6</sup> <https://chat.openai.com/>

级，系统通过将以上信息嵌入人工精心撰写的文本中以生成评价信息。评价信息还包含了对于当前营养摄入及价格的分析及推荐。

#### 目标拆解：

- 1、准确识别食物并计算总价格
- 2、预测食物的卡路里及其他营养成分
- 3、根据食品种类、营养成分和价格等给出评价和膳食推荐

#### 完成情况：

我们顺利完成了最初设定的所有目标，并且将其整合起来形成一个完备的系统，其中食品识别准确率高达 95% 以上，卡路里检测(高中低)结果也十分合理，膳食评价及推荐在结合强大的 ChatGPT 情况下效果颇优。

#### 项目成果(项目技术)总结：

**1、细粒度食物识别模块：** 对模型进行训练使其具有识别分析食物的能力。为了提高模型的鲁棒性和可落地性，我们需要在多个食物数据集上进行训练，并且需要能够实现视频流实时检测食物。经过了前期探索实践，我们自主构建的山大曦园食堂食物数据集<sup>7</sup>的食物检测模块训练效果并不好，详情见附录。由于 YOLO 系列模型在目标检测领域有着准确率高、速度快的特点，我们最终选择使用 YOLOv5 作为模型训练的骨干网络，利用 YOLOv5 进行细粒度的食物识别。另外，通过分析我们自行构建的食物数据集，我们发现中国食物有着菜品多样、成分复杂等特点，也就是说中国菜有很多种菜的种类，并且每一种菜可能由十余种成分炒制而成，这对于现有的深度学习网络来说颇有挑战。所以，为了进一步简化问题，我们选择了较为简洁的日本菜及西餐进行识别。我们使用的数据集为 FOOD100 数据集，该数据集包含 100 种食物照片，每张食物照片都有一个表示食物在照片中的位置的边框。这个数据集中的大多数食物类别都是日本的流行食物。

<sup>7</sup> <https://universe.roboflow.com/juntong-ni-ib1wo/food-detection-c6syr>

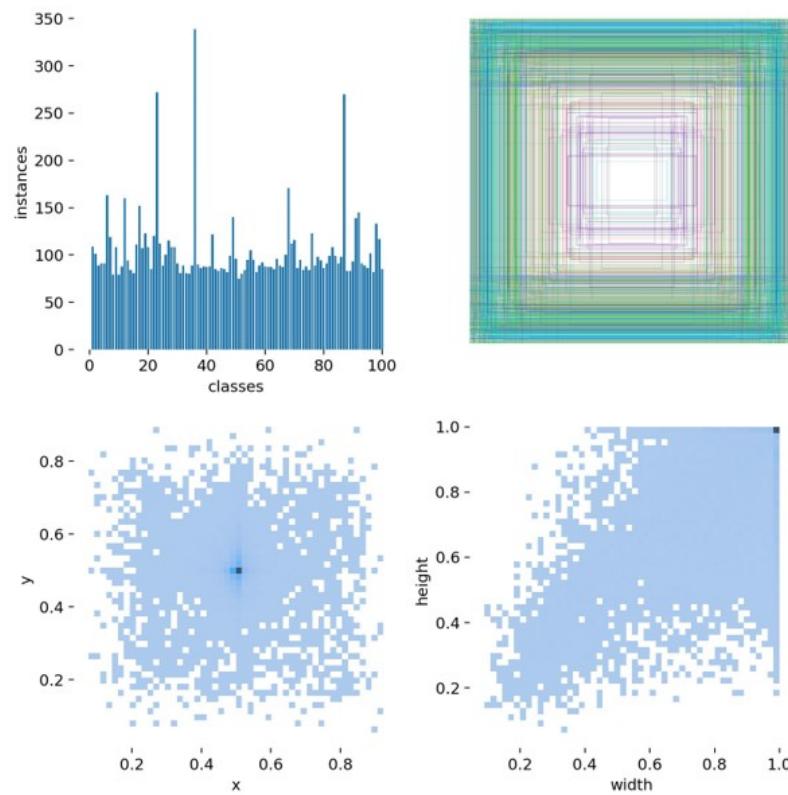


图 1 FOOD 100 数据集统计信息

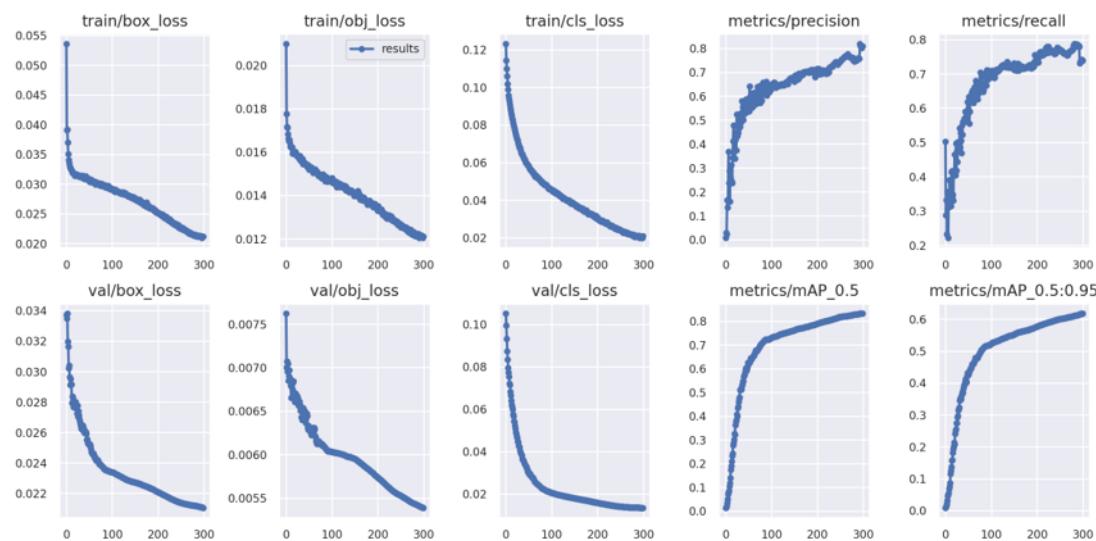


图 2 YOLOv5 训练过程记录

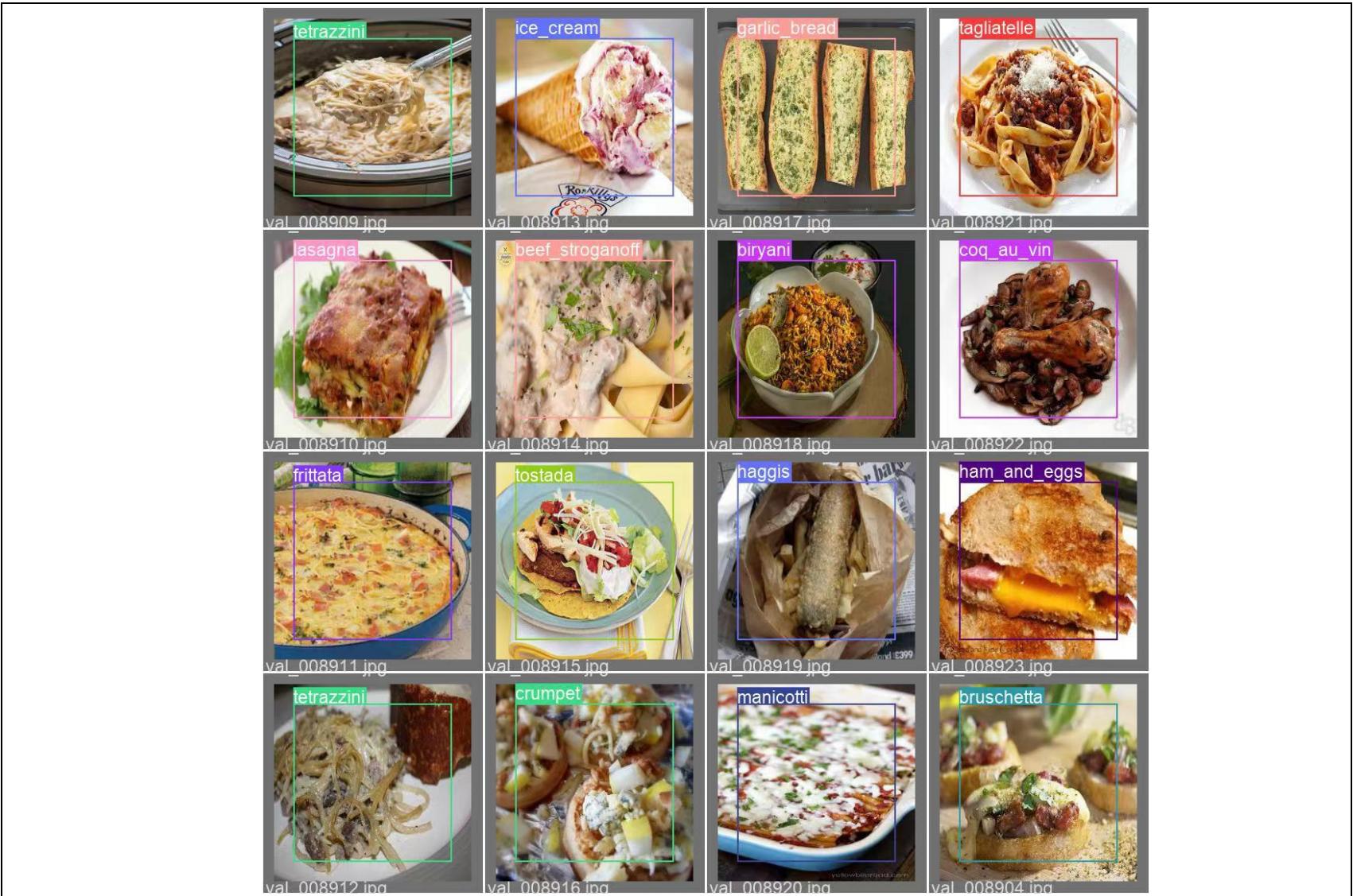


图 3 YOLOv5 训练实际准确度展示

如图 1 所示，第一张图表示各类标签的数量统计，除个别标签数量偏多外，其余标签分布较均匀，第三张图描述了框的中心点分布，第四张图描述了框的宽和高分布。总体而言，我们选用的数据集质量是非常不错的。如图 2 所示，我们将 YOLOv5 在 FOOD100 数据集上训练了 300 个 epochs，`box_loss`、`obj_loss`、`cls_loss`、`precision`、`recall`、`mAP_0.5` 和 `mAP_0.95` 指标均在前 50 个 epochs 快速收敛，并且在后 250 个 epochs 缓慢收敛，最后趋于收敛。如图 3 所示，图中有脆皮奶、冰淇淋、大葱面包、松脆饼、蛋饼等食物。

**2、食物价格匹配模块：**对于所选食物进行价格显示并自动结账。我们需要将训练好的模型应用于食堂智慧结账这一场景中，所以需要具备标价功能。但考虑到人工智能方法对于中国菜品进行识别的准

确率较低，无法直接应用于现实结账情景，我们考虑不在中国菜品上应用我们的方法。我们选用了包含西方菜或者日本菜的数据集进行训练，也就是 FOOD100 数据集，这类菜品比较容易识别，所以识别准确率较高。

在食物细粒度检测的基础上，我们实现了智慧结账系统，在完成食物检测后，模型会根据预先设置的外部价格知识输出最终的价格并进行收费结账。根据实际应用场景，我们的结账系统是能够在视频流中实时检测食物进行结账的，这样才符合我们智慧结账的初衷。



图 4 类别与价格相互整合

我们已经基本将结账系统与食物识别模块进行了整合，由图 4 可见，食物识别的结果还包含了价格信息，且对于存在多个食物的帧，会对每个识别结果的价格进行求和，预期与支付宝等支付方式进行绑定，然后进行自动结账。

**3、食物卡路里分析模块：** 估计每种食物中所含的卡路里含量，然后对比用户输入的可接收的阈值来输出卡路里是高还是低，以提醒用户，保障人体健康。卡路里估算这一功能可以作为食物推荐的依据，为用户提供健康的膳食搭配。我们在进行调研后确定了食物卡路里估算的方法，我们还需要进行图像的语义分割。由于细粒度食物识别是一种目标检测任务，最后为我们呈现的是一个个图片上的选框。因为选框的面积并不能代表食物真实的体积，所以这样的选框不利于我们进行进一步的分析计算。通过像素级图像语义分割，我们可以把不同的食物像素级的分隔开，提高系统的准确性，为进一步的分析计算奠定基础。之后我们需要根据平面分割图及其深度信息对食物进行 3D 重建，然后估算体积，再根据专业机构提供的各类食物单位体积的卡路里含量数据计算最终的卡路里总量，然后对比用户输入的卡路里接收阈值，判断食物卡路里含量的高中低等级。但实际操作后，我们认为根据单张图片进行 3D 重建仍然是计算机图形学领域一个亟待解决的难点问题，目前的已知方法效果较差。从图片或视频中了解食物的

营养成分是一个具有挑战性的计算机视觉问题。这一领域的研究仅限于现有的一些数据集，这些数据集缺乏足够的标签来训练具有营养分析能力的模型。而 Nutrition5k 数据集，包括 5000 种真实的食物种类，以及相应的视频流、深度图像、成分和高准确度的营养含量标签。我们通过在该数据集上训练 ResNet 深度卷积神经网络，该模型能够预测复杂的真实食物的热量和营养含量，其准确度甚至有可能超过了专业的营养学家。

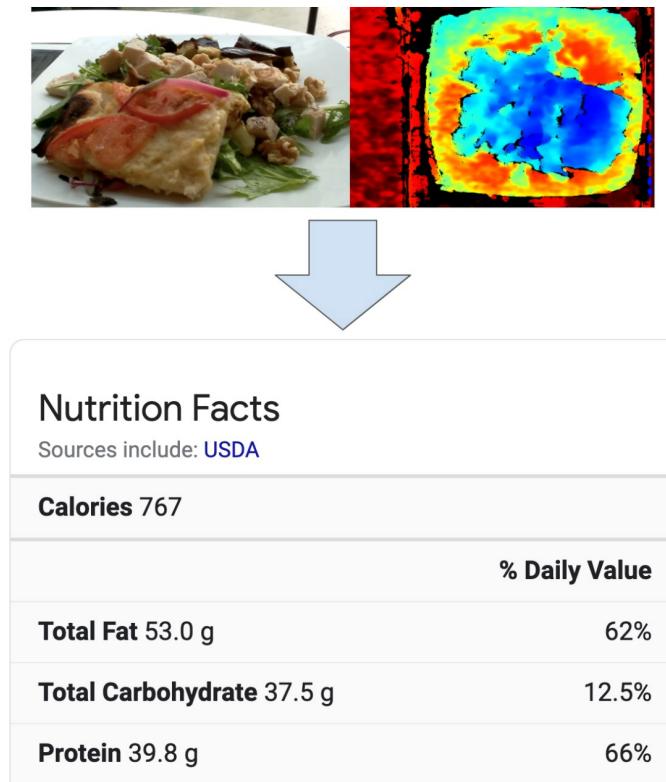


图 5 Nutrition5k 数据集的一个例子

从图 5 中可以看出，Nutrition5K 数据集中包含了一张饭菜的俯视图片及其相应的卡路里、脂肪、碳水化合物、蛋白质等营养成分数值信息。这些信息都是数据集构建者通过专业的测量方法得到的，其准确性可以得到保证。因此，该问题本质上是一个多任务的回归问题，我们在这个问题的基础上训练模型。



图 6 模型训练结果

从图 6 中可以看出，ResNet 在该数据集上训练后（训练时只使用单张俯视图），当其输入为单张照片时，其输出为对应的营养成分含量数值。

**4、营养规划模块：** 我们通过细粒度食物识别模块、食物价格匹配模块以及视频卡路里分析模块得到了食物的种类、价格、卡路里含量等信息。通过综合以上估算预测信息，与人工精心撰写的文本相匹配，生成能够反映用户食物价格高低、营养含量高低的定性评价文本。定性分析的本质是我们根据之前得到的价格信息、卡路里信息等各样信息汇总得出一个评价。这个评价要能够给予用户一定的反馈，告诉用户这顿饭的价格高了还是低了，部分营养指标是否已经高于预期值等，同时给出用户下一顿饭种类的建议，以此来帮助用户进行合理的营养规划。针对该模块，我们设计了两个方案：

第一个是人工撰写。人工针对不同的菜品以及价格、卡路里组合撰写对应的评价推荐文本，这样的人工撰写的文本的优点是更加人性化、更贴合人们的生活实际，但缺点是人工撰写无法覆盖所有的价格和卡路里组合，同时灵活性较差。我们给出了针对不同的菜品以及价格、卡路里组合的人工评价如下

1、菜品：牛排、意大利面；价格：高；卡路里：高

评价：We recommend similar nutritious food, such as chicken breast ,oranges to decrease calorie intake.

2、菜品：番茄、西兰花；价格：低；卡路里：低

评价：We recommend similar nutritious food, such as beef ,fruit salad to increase calorie intake.

3、菜品：意大利面，香肠；价格：中；卡路里：中

评价：We recommend similar nutritious food, such as dumplings ,sliced ham to enrich your meals.

第二个是 ChatGPT 自动生成评价推荐文本。ChatGPT 是由 OpenAI 开发的自然语言处理模型 GPT-3 延伸出的 GPT-3.5 制作的，可以通过大量的数据，来分析和理解输入文本的含义，并判断上下文来生成相应、连贯的回应。而 GPT-3 这个模型也是当前最大的语言处理模型之一，拥有 1750 亿个参数，内容大部分来自网络，包括数十万条维基百科条目、贴文和新闻，让 ChatGPT 能够相当自然地组织语句。

我们利用 ChatGPT 生成的部分评价和推荐如下：

1、 菜品：牛排、意大利面；价格：高；卡路里含量：高

17

我今天吃了牛排和意大利面，价格较高，卡路里含量较高，请你从价格和营养角度进行合理分析并进行食物推荐



从价格的角度来看，牛排和意大利面可能是价格较高的食物。这可能是因为它们使用的原材料或加工方式比较复杂，或者它们的需求量比较大。



从营养的角度来看，牛排和意大利面的卡路里含量较高。这可能是因为它们含有较多的脂肪和/或碳水化合物。尽管这些食物可能有营养价值，但如果您正在尝试降低卡路里摄入量，那么可能需要选择其他食物。

如果您希望在保持营养的同时降低价格和卡路里含量，可以尝试更多的蔬菜和水果，以及谷物，豆类和蛋白质来源，如鸡肉，鱼，豆类或豆制品。这些食物通常价格较低，卡路里含量较低，并且富含营养素。

最后，建议您与医生或营养师进行咨询，以了解您的具体营养需求和最适合您的饮食计划。

2、 菜品：番茄、西兰花；价格：低；卡路里：低

AI

我今天吃了番茄和西兰花，价格较低，卡路里含量低，请你从价格和营养角度进行合理的分析和食物推荐



你今天吃的番茄和西兰花是很好的选择。这两种食物都是健康的选择，因为它们都是低卡路里、低价格的选择。



番茄是一种富含维生素C、维生素A和番茄红素的水果。番茄红素是一种天然抗氧化剂，可以帮助抵抗自由基的损害，保护身体免受疾病的侵袭。番茄还含有钾、纤维和镁等营养成分，可以帮助调节血压、促进肠胃蠕动，并有助于预防心血管疾病。

西兰花是一种富含维生素C、叶酸和钙的蔬菜。叶酸是一种B族维生素，可以促进神经系统的正常发育，帮助人体合成DNA和RNA。西兰花同时也含有大量的纤维、钾、镁和硒，有助于促进肠胃蠕动、调节血压，并有助于预防心血管疾病。

此外，你还可以尝试吃一些其他的低卡路里、低价格的健康食物，如豆类、蔬菜、水果、全麦食品等。这些食物都是富含营养的选择，可以帮助你保持身体健康。

3、 菜品：意大利面，香肠；价格：中；卡路里：中

AI

我今天吃了意大利面和香肠，价格合适，卡路里含量中，请你从价格和营养角度进行合理的分析和食物推荐



意大利面和香肠是一种常见的选择，但它们的营养价值和卡路里含量可能较高。意大利面通常是由小麦粉和水制成的，它含有大量的碳水化合物和蛋白质，但可能不含有太多的营养成分。香肠通常是由猪肉、盐、调味料和其他香料制成的，它含有大量的脂肪和蛋白质，但可能不含有太多的营养成分。



如果你想要更加健康的食物选择，可以考虑尝试一些富含营养的食物，如蔬菜、水果、豆类、全麦食品等。这些食物通常富含纤维、维生素和矿物质，并且卡路里含量较低。此外，你也可以尝试一些健康的谷物，如大米、黑豆、燕麦等，这些食物通常富含蛋白质、膳食纤维和抗氧化剂，可以帮助你保持身体健康。

通过观察人工撰写的营养规划文本和 ChatGPT 生成的文本，我们可以观察到，人工撰写的文本内容的丰富程度远不如 ChatGPT 生成的文本。且 ChatGPT 生成的文本不仅包含了对当前食物的营养分析，还包括了一部分未来的食物推荐。虽然 ChatGPT 目前也存在一定问题，但我们认为其应用前景非常广阔，值得期待。

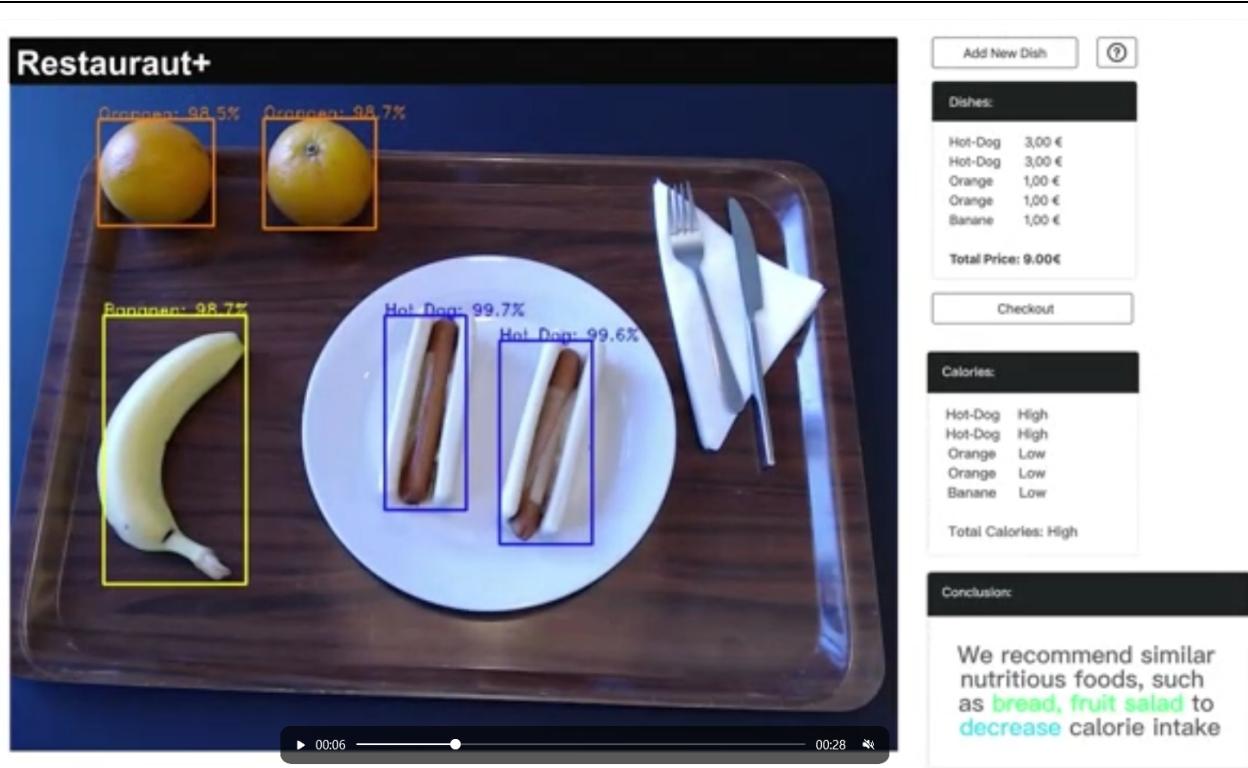
## 项目结果与分析：

### 项目结果：

我们最终完成了一个集细粒度食物识别、卡路里估计、食品推荐于一体的膳食推荐系统。在托盘整体放入摄像头不到 3s 内，我们的模型就已给出了食物识别、卡路里估计和食品推荐的结果，可见实时性非常好。

下面是最终结果展示（实际效果是实时检测，这里不方便嵌入视频，故以图片形式展示），一共展示了四个例子。在每一个例子中左侧是待检测的食物，每类食物上会打上一个框及对应名称。右侧第一栏是食物名称及对应价格和总价格，第二栏是食物对应的卡路里(高中低)，第三栏是我们模型给出的人工撰写的推荐。

### 第一个例子：



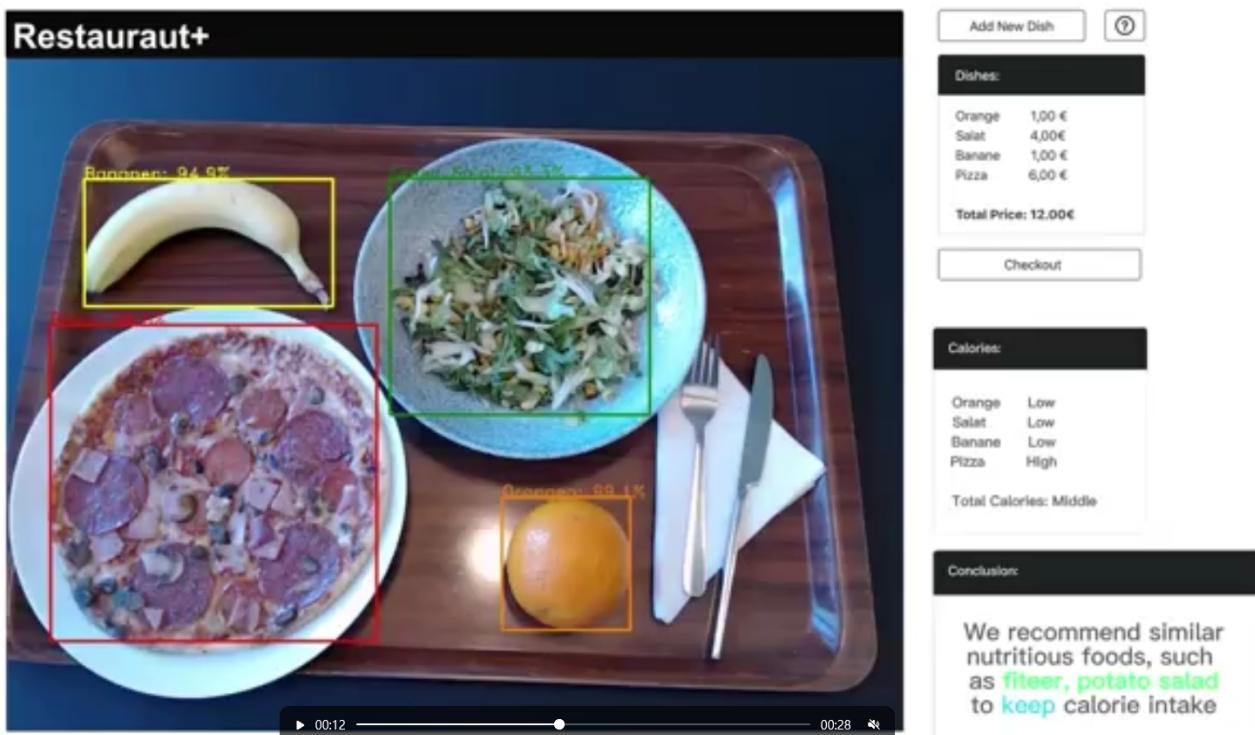
成功识别出两个橘子，一根香蕉，两根热狗。橘子和香蕉是低热量食物，热狗是高热量食物。模型分析摄入的总热量是高的，故推荐之后食用面包和水果沙拉来减少摄入的卡路里。

#### 在 ChatGPT 上生成建议：

输入：我今天吃了两个橘子，一根香蕉，两根热狗，卡路里含量高，请你从价格和营养角度进行合理的分析和食物推荐。

输出：都是很好的选择。橘子富含维生素 C，可以帮助改善免疫力，预防感冒。橘子还含有钾、镁和纤维等营养成分，有助于调节血压、促进肠胃蠕动，并有助于预防心血管疾病。香蕉是一种富含钾的水果，可以帮助调节血压，促进肠胃蠕动。香蕉还含有大量的纤维，可以帮助降低胆固醇，预防心血管疾病。但是，热狗通常是由加工肉类、盐、调味料和其他香料制成的，它们含有大量的脂肪、盐和碳水化合物，而且卡路里含量较高。如果你想要更加健康的食物选择，可以尝试一些低脂肪、低盐的食物，如蔬菜、水果、豆类、全麦食品等。这些食物通常富含纤维、维生素和矿物质，并且卡路里含量较低。此外，你也可以尝试一些健康的谷物，如大米、黑豆、燕麦等，这些食物通常富含蛋白质、膳食纤维和抗氧化剂，可以帮助你保持身体健康。

第二个例子：



成功识别出一个橘子，一根香蕉，一盘披萨和一盘沙拉。橘子、香蕉和沙拉是低热量食物，披萨是高热量食物。模型分析摄入的总热量是中的，故推荐之后食用土豆沙拉来保持摄入的卡路里，并丰富自己的膳食。

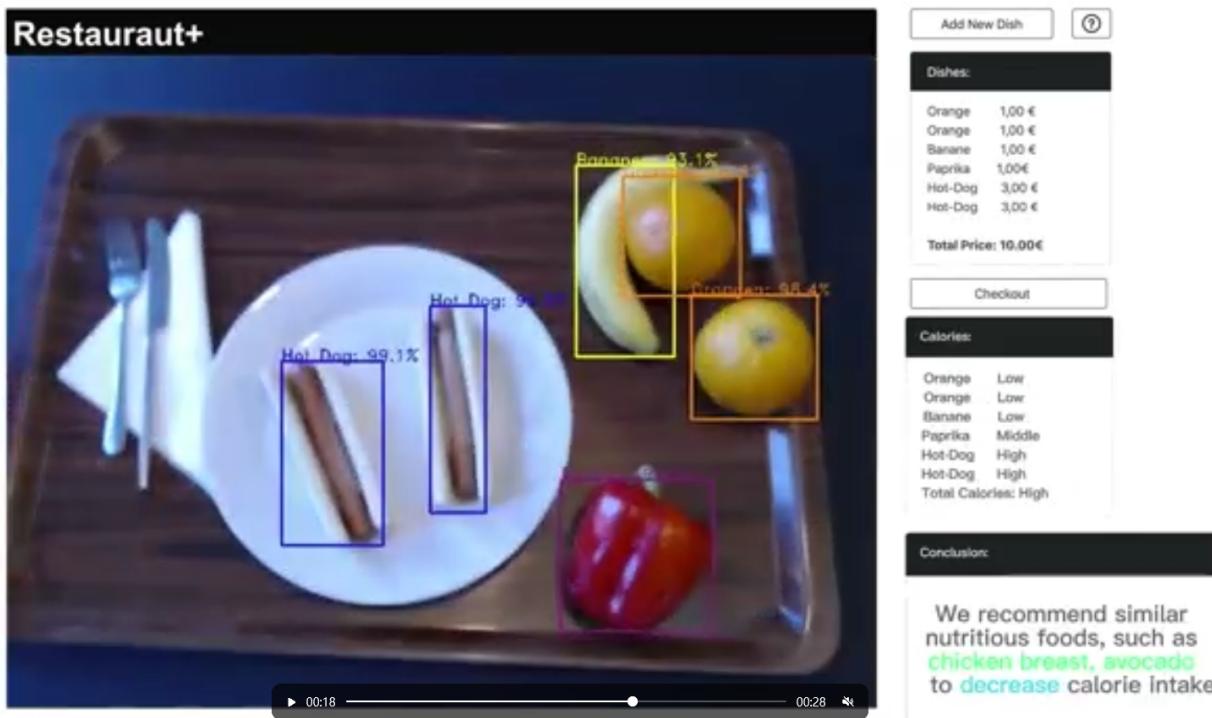
在 ChatGPT 上生成建议：

**输入：**我今天吃了一个橘子，一根香蕉，一盘披萨和一盘沙拉，卡路里含量中，请你从价格和营养角度进行合理的分析和食物推荐。

**输出：**你今天吃的橘子、香蕉和沙拉都是富含营养的食物，可以为你的身体提供大量的维生素和矿物质。橘子是一种富含维生素 C 的水果，可以帮助提高免疫力，促进皮肤健康。香蕉是一种富含镁、钾和钙的水果，可以帮助调节血压、促进肠胃蠕动。沙拉是一种富含纤维、维生素和矿物质的蔬菜，可以帮助促进肠胃蠕动、调节血压，并有助于预防心血管疾病。然而，你今天吃的披萨可能对你的身体不太友好。披萨通常是由高碳水化合物、高脂肪和高蛋白质制成的，卡路里含量较高，如果过多地食用可能会对你的健康造成不良影响。如果你想要更加健康的食物选择，可以考虑尝试一些富含营养的食物，如蔬菜、水果、豆类、全麦食品等。这些食物通常富含纤维、维生素和矿物质，并且卡路里含量较低。此

外，你也可以尝试一些健康的谷物，如大米、黑豆、燕麦等，这些食物通常富含蛋白质

第三个例子：



成功识别出两个橘子，一根香蕉，两根热狗，一个番茄。橘子、香蕉和番茄是低热量食物，热狗是高热量食物。模型分析摄入的总热量是高的，故推荐之后食用鸡胸肉和牛油果来减少摄入的卡路里。

在 ChatGPT 上生成建议：

**输入：**我今天吃了两个橘子，一根香蕉，两根热狗，一个番茄，卡路里含量高，请你从价格和营养角度进行合理的分析和食物推荐

**输出：**橘子、香蕉和番茄是很好的水果选择，它们都富含维生素 C、维生素 A 和纤维等营养成分。但是，热狗通常是由猪肉、小麦粉和调味料制成的，它含有大量的脂肪、碳水化合物和盐，卡路里含量相对较高。如果你想要更加健康的食物选择，可以考虑尝试一些富含营养的食物，如蔬菜、水果、豆类、全麦食品等。这些食物通常富含纤维、维生素和矿物质，并且卡路里含量较低。此外，你也可以尝试一些健康的谷物，如大米、黑豆、燕麦等，这些食物通常富含蛋白质、膳食纤维和抗氧化剂，可以帮助你保持身体健康。

第四个例子：

The screenshot shows a food recognition application interface. On the left, there is a photograph of a meal consisting of spaghetti, two oranges, and a banana. Overlaid on the image are several colored bounding boxes with confidence percentages: a pink box on the spaghetti (98.5%), a yellow box on the banana (99.3%), and an orange box on each of the two oranges (98.5% and 99.3%). The right side of the screen contains a user interface with the following sections:

- Dishes:**

Orange	1,00 €
Banane	1,00 €
Spaghetti	6,00 €
Orange	1,00 €

Total Price: 9.00€
- Checkout**
- Calories:**

Orange	Low
Banane	Low
Spaghetti	Middle
Orange	Low

Total Calories: Low
- Conclusion:**

We recommend similar nutritious foods, such as [lasagna](#), [durian](#) to increase calorie intake.

成功识别出两个橘子，一根香蕉，一盘意大利面。橘子和香蕉是低热量食物，意大利面是中热量食物。模型分析摄入的总热量是中的，故推荐之后食用宽面和榴莲来保持摄入的卡路里，并丰富自己的膳食。

在 ChatGPT 上生成建议：

输入：我今天吃了两个橘子，一根香蕉，一盘意大利面，卡路里含中，请你从价格和营养角度进行合理的分析和食物推荐

输出：橘子和香蕉是一种很好的水果选择，因为它们都富含维生素 C 和其他营养成分。橘子是一种富含维生素 C、钾和纤维的水果，可以帮助调节血压、促进肠胃蠕动，并有助于预防心血管疾病。香蕉是一种富含钾、纤维和维生素 B6 的水果，可以帮助调节肠胃蠕动、促进肌肉健康，并有助于预防心血管疾病。意大利面是一种常见的选择，但它的营养价值和卡路里含量可能较高。意大利面通常是由小麦粉和水制成的，它含有大量的碳水化合物和蛋白质，但可能不含有太多的营养成分。如果你想要更加健康的食物选择，可以考虑尝试一些富含营养的食物，如蔬菜、豆类、全麦食品等。这些食物通常富含纤维、维生素和矿物质，并且卡路里含量较低。此外，你也可以尝试一些健康的谷物，如大米、黑豆、燕麦等，这些食物通常富含蛋白质、膳食纤维和抗氧化剂，可以帮助你保持身体健康。

## 项目进行时出现的问题及解决方案：

**问题 1：**在结账系统中，由于食物识别不能达到几乎 100%的准确率，而收银这件事又是非常严谨的，需要丝毫不能出错。所以根据食物识别的结果进行收账是不现实的，需要另想办法达到基本全对的准确率。

**解决：**根据老师的指导与团队内的讨论，可以根据餐盘的颜色来对应不同的价格，因为普通的颜色识别是最基本的任务，在如今的深度学习模型上，基本可以达到百分比的准确率。或者根据每个食物对应的二维码进行扫描等技术达到超高精度。我们的解决方法是不采用复杂的中国菜数据集，该用花样简单的日本菜或西方菜数据集。

**问题 2：**为了增强项目的实际落地性，我们尝试了制作属于自己的数据集，但是事情远远没有想象的简单。比如由于食堂某窗口的人流量不固定，采集的图片数量不够；由于食堂的饭菜在餐盘的位置并不固定，会有重叠交叉部分；由于食堂的中国菜可混搭性很强，比如鸡蛋基本可以和任何菜混搭，显然不能一样的一个标签，否则导致标签数量过多，因此造成标签不明确的问题。

**解决：**对于采集的图片不够的问题，我们采用了图像增广的办法先进行一次离线版本的扩充数据集，可以将数据集扩充两三倍，基本可以达到目标要求。

对于食物位置重叠交叉部分，我们决定弃置此张照片，因为质量较差的图片反而会降低训练效果；对于标签不明确问题，还没有很好的解决办法，只能尽量将混搭菜减少，相似的进行统一。

**问题 3：**目前卡路里的估算方法要么精度太差，要么没有开源，导致最终的卡路里估算结果不太可信。

**解决：**根据老师的指导和团队内的讨论，可以将具体的卡路里数字结果转换成分类问题，及根据用户预设的可接收阈值，系统经过卡路里估算后会输出所含的卡路里高中低等级，这样的设置增大了可信度与合理性。

## 分析与体会：

总体而言，我们项目的完成度非常高，不仅实现了各部分的功能还将其组合成一个整体，实用性和落地性好，最关键的是实时性强。

**1、 细粒度食物识别部分:**我们使用了在物体检测任务中表现良好的 YOLOv5 模型，在日本菜数据集

FOOD100 上进行训练，在测试集上表现优异准确率高。但也存在一定的缺陷，比如食物类型不够，在检测其他菜系时可能并不是很准确，若要落地则需要针对相应的数据集进行训练。我们采用的食品价格是参考 NUMBEO 食品价格数据库中的信息。

- 2、卡路里估计部分：由于在这一部分我们很难使模型对卡路里的计算误差在一个合理的范围内，故我们转换思路，只估计食物卡路里的大致范围(高中低来衡量)
- 3、食品推荐部分：这一部分我们既尝试使用人工精心编写的评价及建议，也尝试了最近大火的 ChatGPT 模型来生成评价和建议。两者互为补充，让我们的食品推荐更加丰富。

### 总结与反思：

- 1、首先对于项目的展开，不能着急定题，应该对于所选题目做好充分调研之后再确定，比如找好各个模块需要参考的资料，而不是大致制定计划就行，否则就会出现项目卡壳的阶段，找不到合适的参考资料而停滞一段时间。
- 2、然后是制作数据集的困难，当时为了增强项目的实际落地性，我们计划制作山大食堂数据集，当时预期应该没什么困难。但是在收集阶段时就出现所选窗口的人流量不固定，收集不到足够的数据集；在数据标注阶段发现食物之间出现重叠交叉部分且由于食堂里中国菜的可搭配性强，导致标签不明确的种种问题。深深体会到了前人制作优秀数据集的困难。
- 3、对于项目的设计太过于繁多沉重，对于卡路里估算和推荐系统而言，本身就算复杂困难的项目，而这又被我们希望集成到系统之中，这就造成项目的困难，工作量大。按照老师的建议，我们对相关模块进行简化，同时提高合理性和可信度。所以项目的设计应该考虑好工作量以及现实的问题，是否合理可信等。

### 参考文献与资源：

#### 1、YOLO

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

#### 2、ResNet

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and

*pattern recognition* (pp. 770-778).

### 3、Food100 Dataset

Matsuda, Y., Hoashi, H., & Yanai, K. (2012, July). Recognition of multiple-food images by detecting candidate regions. In *2012 IEEE International Conference on Multimedia and Expo* (pp. 25-30). IEEE.

### 4、Nutrition5k Dataset

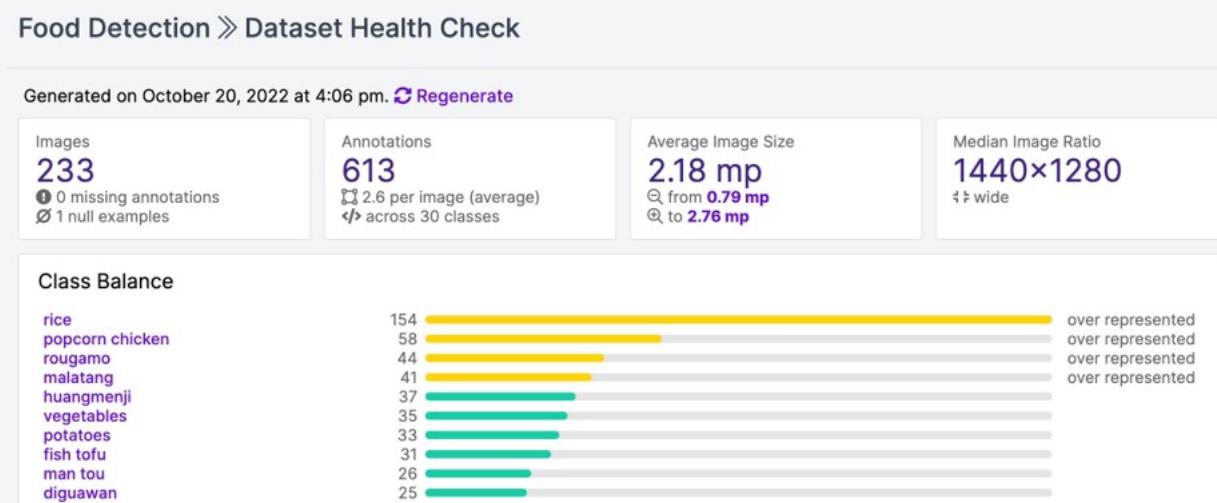
Thames, Q., Karpur, A., Norris, W., Xia, F., Panait, L., Weyand, T., & Sim, J. (2021). Nutrition5k: Towards automatic nutritional understanding of generic food. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8903-8911).

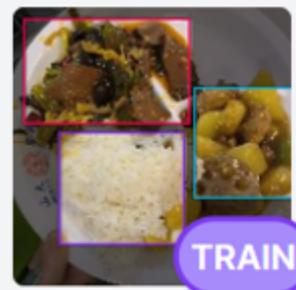
## 附录：

### 1、自主构建的曦园食堂的数据集

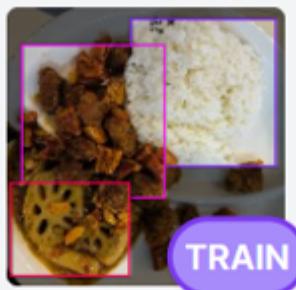
为了让我们的项目能够落地应用，我们在山东大学曦园食堂进行了为期两星期的食物数据收集活动，经过首次图像增广后，收集了 500 张左右的食物图片。之后我们三人分工对收集到的食物图片进行标注，标注格式参考 coco 数据集。完成后使用之前预训练的 **yolov5** 模型在该数据集上进行训练。效果良好。并同时实现了食堂智慧结账系统，预期结合校园卡，微信支付等支付方式实现智慧结账功能。

其中收集并标注数据集的工作比预想中困难不少，由于团队只有三人，想要收集高质量的数据集需要花费大量时间。且所选的食堂窗口，人流量不固定，所能收集的图像较少，且质量高低不齐，会出现饭菜之间混合在一起问题。在标注数据时，各个菜的数量不一致，出现明显区别，有点甚至两极分化，不得已只能放弃数量太少的图片；由于中国菜可大搭配混合性很强，所以在确定标签时也遇到困难。





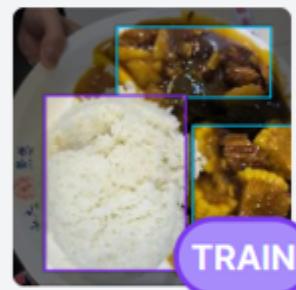
62594ABCB6F8...



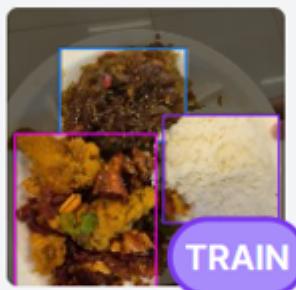
F5933B488CCA...



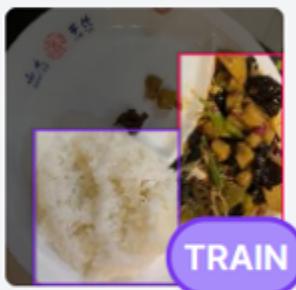
C929F6C8A98B...



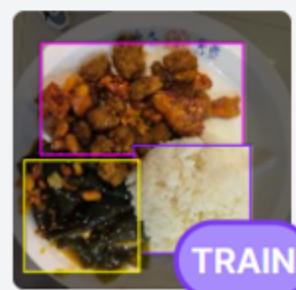
64FDDE46EFF23...



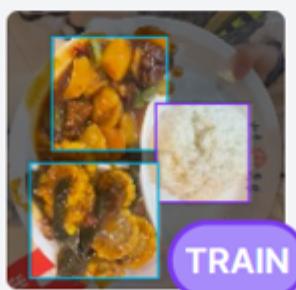
3F90434538E98...



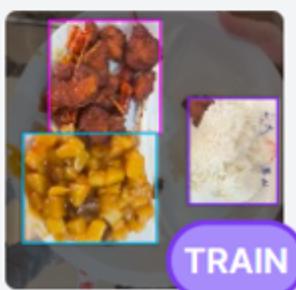
920014FD5D85F...



FEAB49B280413...



D1428237-F14A...



E29A1694-F547...

上面即是我们数据的大体展示，基本图像数量有两百多张，可以进行数据增广来扩大量。由于是食堂的饭菜，每张图像基本有三个目标，基本的类别如图，同时我们舍弃了数量很少的类别，防止降低训练效果。