

DOI: 10.13718/j.cnki.xdzk.2023.04.018

基于图像语义分割的菜品浪费度检测

普京, 祝诗平, 苗宇杰, 唐鑫, 郑权, 黄华

西南大学 工程技术学院, 重庆 400715

摘要: 对菜品浪费进行检测可以帮助餐馆调整菜单结构, 做到从源头上减少浪费。提出一种基于语义分割技术的菜品浪费度检测方法, 使用 Swin Transformer 和 UperNet 作为分割框架, 并在 UperNet 的基础上进行了改进, 即在 UperNet 的双线性插值层之前增加了一层转置卷积, 改进的模型在测试集上的菜品图像分割平均交并比可以达到 93.30%; 同时根据菜品食用前后分割面积比的统计信息制定了浪费度等级, 将菜品浪费程度划分为严重浪费、很浪费、一般浪费、轻度浪费和无浪费 5 个等级, 并在测试集 1 782 个菜品样本上进行了检测, 浪费度等级的识别准确率可以达到 95.12%。所给出的方法可望用于餐饮浪费的检测和管理中。

关键词: 菜品; 图像; 语义分割; 深度学习; 浪费度等级

中图分类号: TP399

文献标志码: A

文章编号: 1673-9868(2023)04-0189-12

开放科学(资源服务)标识码(OSID):



Detection of the Degree of Dish Waste Based on Image Semantic Segmentation

PU Jing, ZHU Shiping, MIAO Yujie,
TANG Xin, ZHENG Quan, HUANG Hua*College of Engineering and Technology, Southwest University, Chongqing 400715, China*

Abstract: Detecting the waste of dishes can help restaurants to adjust the menu structure and reduce the waste from the source. This study proposed a method to detect the degree of dish waste based on semantic segmentation technology. Using Swin Transformer and UperNet as the segmentation framework, the UperNet was improved by adding a transpose convolution layer before the bilinear interpolation layer of the UperNet. Finally, the average intersection over union of the improved model on the test set for image segmentation can reach 93.30%. In addition, according to the statistical information of the segmented area

收稿日期: 2022-04-29

基金项目: 教育部科技发展中心产学研创新基金德州专项(2021DZ005)。

作者简介: 普京, 硕士研究生, 主要从事智能检测与控制研究。

通信作者: 祝诗平, 教授, 博士研究生导师。

ratio of dishes before and after consumption in the training and verification set, this paper also formulated the category of the waste, which divided the waste degree of dishes into five grades: serious waste, very waste, general waste, mild waste, and no waste. Of 1 782 dish samples tested, the accuracy rate of waste degree identification can reach 95.12%. The method presented in this paper is expected to be used in the detection and management of food waste in the future.

Key words: dishes; image; semantic segmentation; deep learning; degree of waste

粮食从农田到餐桌,需要经过种植、收获、运输和储藏等诸多环节,每一个环节都蕴含着劳动者的心血和智慧,节约粮食,我们每个人都责无旁贷。然而,随着人们生活水平的提高和生活节奏的加快,食物浪费现象也越来越严重。李贺等^[1]对我国餐厅食物浪费现状进行了调查研究,发现消费者餐饮浪费的主要原因为菜品不合胃口以及菜品份量较大等。据统计,近年来我国的城市餐厨垃圾产量占生活垃圾产量的 55.86%,且年均增速达 10%以上^[2-3],常规的餐厨垃圾处理方法会对环境造成较大的危害^[4]。此外,食物浪费现象也和粮食安全问题有着密切的关联,据联合国粮农组织统计,全球餐厨垃圾产量每年约为 13 亿 t,约占年食物制造总量的 1/3,这一数量足以支撑世界 12%的人口摆脱饥饿^[5],因此,从源头消减和抑制餐饮浪费十分必要。

对餐桌上的菜品浪费情况进行监测,可以帮助餐馆淘汰掉浪费程度较高的菜品,但首先需要解决菜品种类的识别问题。Chen 等^[6]从网络上收集并整理了中餐菜品数据集 VireoFood-172,提出了菜品种类识别以及食谱检索模型。针对菜品图像识别过程中存在的不同类别菜品间的相似度较高(类间距离小),而有些相同种类的菜品其颜色、形状变化较大(类间距离大)的问题,一些学者对传统的识别模型进行了改进,引进了基于细粒度的识别模型以及集成学习方法,使菜品种类的识别准确率得到提升^[7-9]。比种类识别更进一步,一些学者研究了菜品识别中的目标检测问题^[10, 11],可以定位出菜品图像边界框区域。比菜品目标检测任务更进一步的就是菜品图像的分割问题,Poply 等^[12]采用目标检测框架和语义分割模型相结合的方式对菜品图像进行了分割,并由先验知识估计出菜品的体积和质量,从而进一步估算出菜品所含热量。Okamoto 等^[13]从网络上收集菜品所含的热量信息,采用有监督学习的方式,在 DeepLab V3+^[14]的基础上进行了改进,融合了菜品图像分割和菜品所含热量估计的任务。

很多学者都对菜品图像的种类识别、目标检测以及分割问题进行了探讨,然而针对菜品浪费度的检测问题目前还鲜有报道。对菜品浪费情况进行检测,一种可能的方法就是估计出菜品的质量或体积,Hippocrate 等^[15]借助手机相机和餐具(如筷子)来估计食物的质量,进而估计食物所含的热量,但该方法是通过测算碗的体积,并提供特定食物的密度信息来估计食物质量,也就是说该方法只能估计盛满食物状态下的质量,而无法估计未盛满时的质量,因此很难应用于食物浪费情况的检测中。Chen 等^[16]通过 SIFT^[17]特征提取的方法来识别菜品种类,并用深度相机来估计食物体积,但不同食物对红外线的折射与反射效果不同,深度相机发出的红外线很容易受到影响,导致测量结果不准确。Hu 等^[18]采用了三维点云技术对食物的体积进行估计,但该方法需要采集到多个视觉下的图像并对场景进行三维重构,对图像采集的条件也较为苛刻,所采集食物的形状也要求较为规则,因此也很难应用于菜品浪费检测中。

针对以上问题,本文提出了通过语义分割技术来检测菜品的浪费程度,根据菜品食用前后分割得到的像素面积之比来衡量菜品的浪费程度,并根据训练集和验证集上菜品食用前后分割得到的像素面积比的统计信息,制定相应的浪费度等级。在图像语义分割领域,自全卷积神经网络 FCN^[19]提出以来,深度学习方法在图像分割领域的应用变得越来越流行,后续也出现了一系列优秀的基于卷积神经网络的分割结构^[14, 20-21],Liu 等^[22]提出了基于移动窗口的自注意力模型 Swin Transformer,该模型在计算机视

觉的下游任务(如目标检测和语义分割)上取得了优异的成就,在大数据集上的实验结果表明其性能优于卷积神经网络结构,因此,本文选用 Swin Transformer 结构作为主干特征提取网络,解码器部分则在 UperNet^[23]的基础上进行了改进,在双线性插值层之前增加了一层转置卷积操作,进一步提高了菜品图像的分割效果。

1 材料与方法

1.1 数据采集及数据预处理

1.1.1 数据采集

本实验所采集的数据为某高校食堂提供的菜品,实验使用 1 部红米 note 9 pro 手机和 1 个手机固定架来完成图像采集,手机分辨率为 $2\,400\times1\,080$ 像素,采集过程中将手机固定在菜品上方,记录下餐盘里的菜品从有到无的变化过程(图 1)。最终采集到的图片数量共为 3 765 张。每一张图片中包含 3 到 5 个菜品样本,总的菜品样本数量为 12 200 个,总的菜品类别数为 25 类(图 2)。将 25 类菜品按 1 到 25 进行编号,每一类菜品的数量分布情况如图 3。

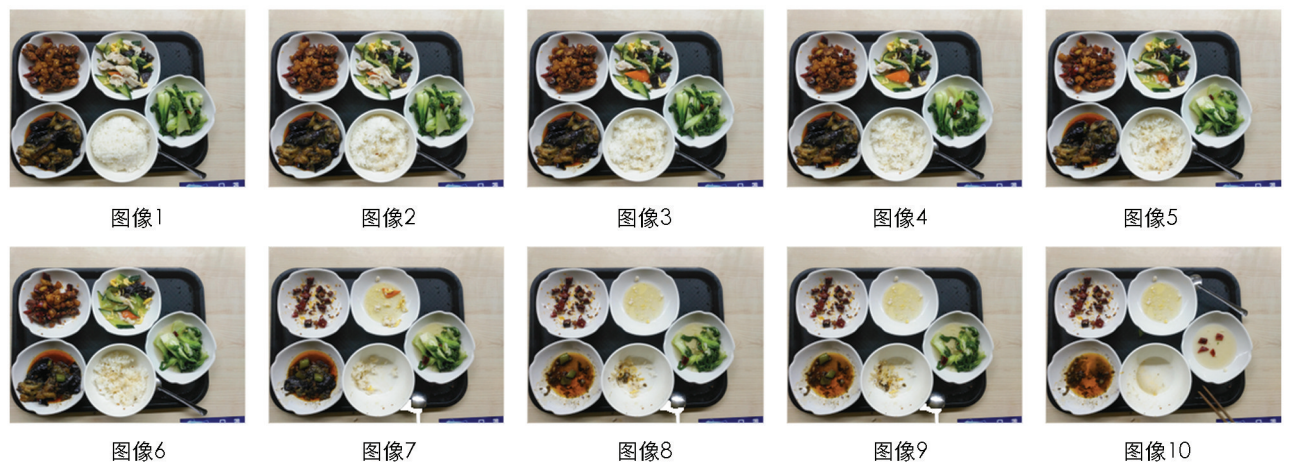


图 1 菜品图像变化过程

1.1.2 数据标注

使用 Labelme 和 EIsseg 软件对采集好的图像进行人工标注,生成对应的分割掩码信息。标注过程中只标注每一类菜品的可食用区域,而不宜食用的菜品如花椒、辣椒、油等则标注为背景区域(图 4)。

1.1.3 数据集划分

将 3 765 张数据划分为训练集、验证集和测试集,训练集用于模型训练,一共 2 440 张,包含 8 163 个菜品样本;验证集数据一共 698 张,包含 2 255 个菜品样本,该部分数据用于判断模型的收敛性以及模型的训练效果,并控制模型达到一定条件后终止训练;测试集数据有 627 张,菜品样本数量为 1 782 个,用于检验模型的泛化能力和最终效果。测试集部分的数据单独进行采样,训练集、验证集的数据则是在所有图片打乱顺序后进行随机划分的。

1.1.4 数据增强

为了增强模型的鲁棒性,防止模型过拟合,本实验对训练集的数据进行了增强(图 5),对训练集的图片以及标签图片同时旋转任意角度,来模拟不同角度下的拍摄效果^[24];改变图像的亮度信息,对图像进行色彩抖动,来模拟不同光照条件下的拍摄效果。最终,训练集的数据扩大为原来的 3 倍,共为 7 320 张,包含 24 489 个菜品样本。



图 2 菜品种类

1.2 菜品图像语义分割模型

1.2.1 模型结构

在目标检测、图像分割等图像处理下游任务上 Swin Transformer 都表现出了优异的性能^[22]，相较于传统的卷积运算，Swin Transformer 中基于移动窗口的自注意力方式具有更强的特征提取能力，因此，本文选用了 Swin Transformer^[22]作为特征提取网络。解码器部分则在 UperNet^[23]结构的基础上进行了改进，记为 M-UperNet，模型的整体结构如图 6。

图 6 中 Swin Transformer 部分为主干特征提取网络，其中 Patch Partition 的作用是将输入图像划分成若干个小方格，Linear Embedding 为一次卷积运算，Swin Transformer 模块为一次基于窗口的自注意力运算和一次基于移动窗口的自注意力运算的结合，本文利用了该主干网络 4 个不同尺度的特征层，其特征尺

度分别为 4 倍、8 倍、16 倍和 32 倍下采样. PPM 为金字塔池化模块, 该模块融合了不同尺度的池化操作, 从而能保留图片的全局上下文信息, 提升分割模型的精度. FPN 为图像特征金字塔模块, 该模块将图像的不同特征层进行融合, 利用高级语义信息的同时关注浅层的语义信息.

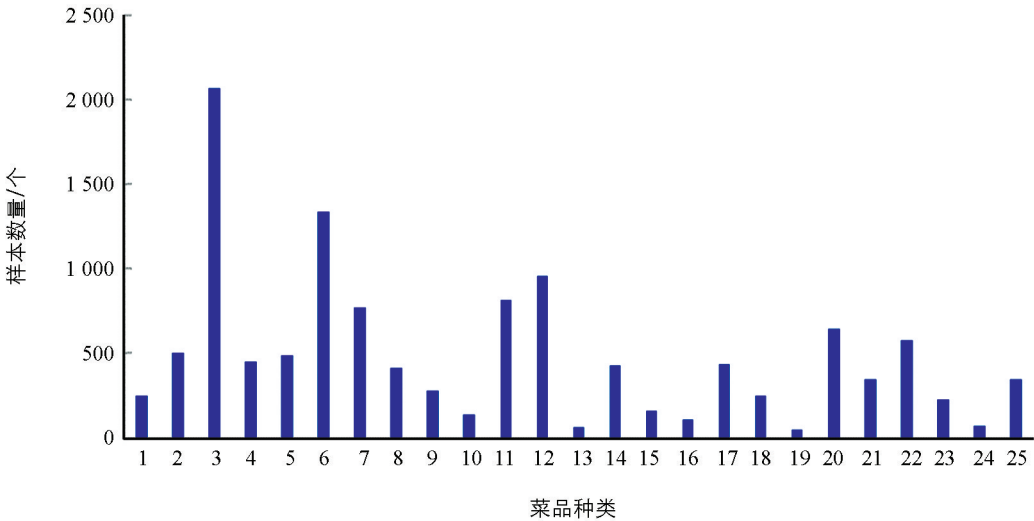


图 3 每一类菜品数量分布情况

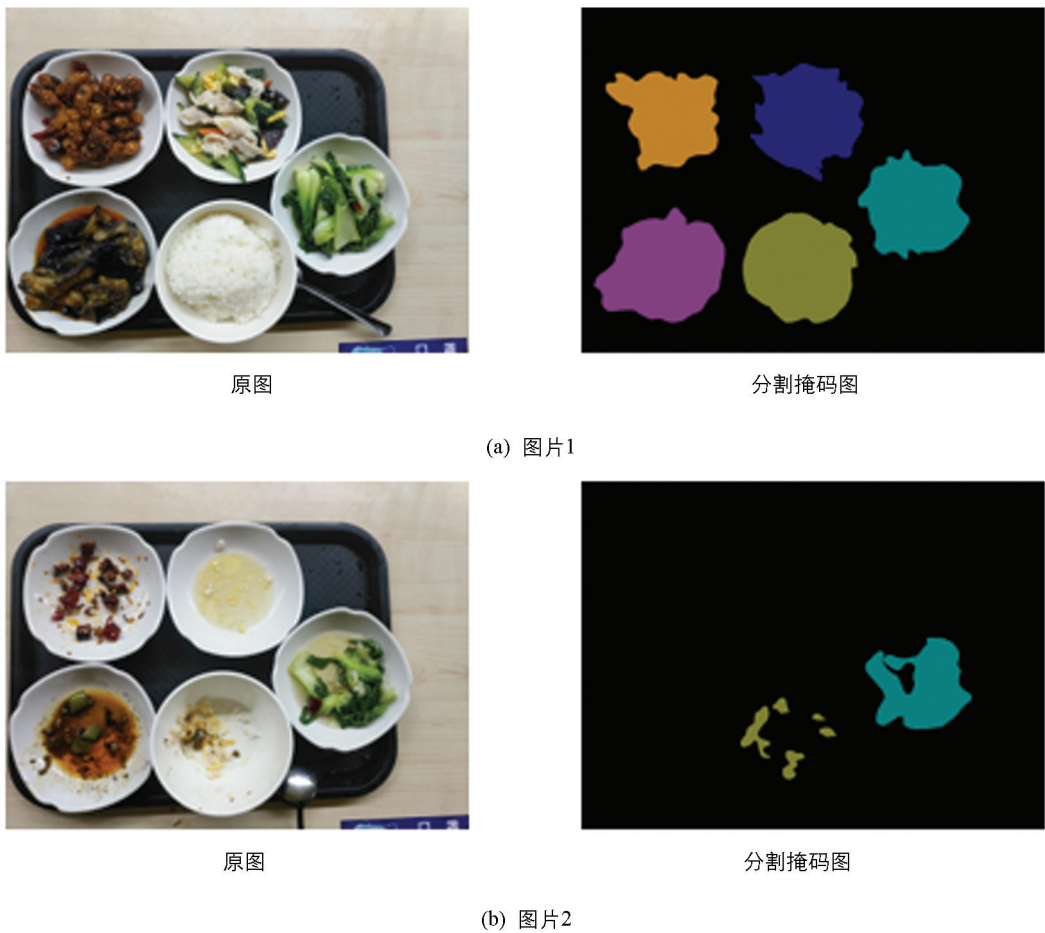


图 4 菜品原图和分割掩码图



图 5 菜品图像数据增强

在 UperNet 结构中, 图像的语义分割结果是在卷积层之后使用一次双线性插值得到的, 不同于连续物体的分割, 菜品浪费度检测的分割过程中会存在以下问题: 当菜品食用到一定程度后, 可食用菜品的区域可能是不连通的, 直接用双线性插值方法还原出和原始图像相同大小的分割掩码时会引入一定的噪声, 从而导致分割边界的模糊, 所以本文在双线性插值之前引入了一层转置卷积^[25]操作, 如图 6 中的转置卷积部分, 卷积核大小为 3, 步长为 2. 本文中输入图像的大小为 224×224 像素, 卷积层得到的特征图大小为 56×56 像素, 经过转置卷积层, 得到的特征图大小变为 113×113 像素, 最后再经过一次双线性插值运算还原出和输入图像相同大小的掩码. 通过加入转置卷积层, 可以引入可训练的参数, 从而减少信息的损失, 使分割结果更为精细. 本文将改进后的解码器结构记为 M-UperNet.

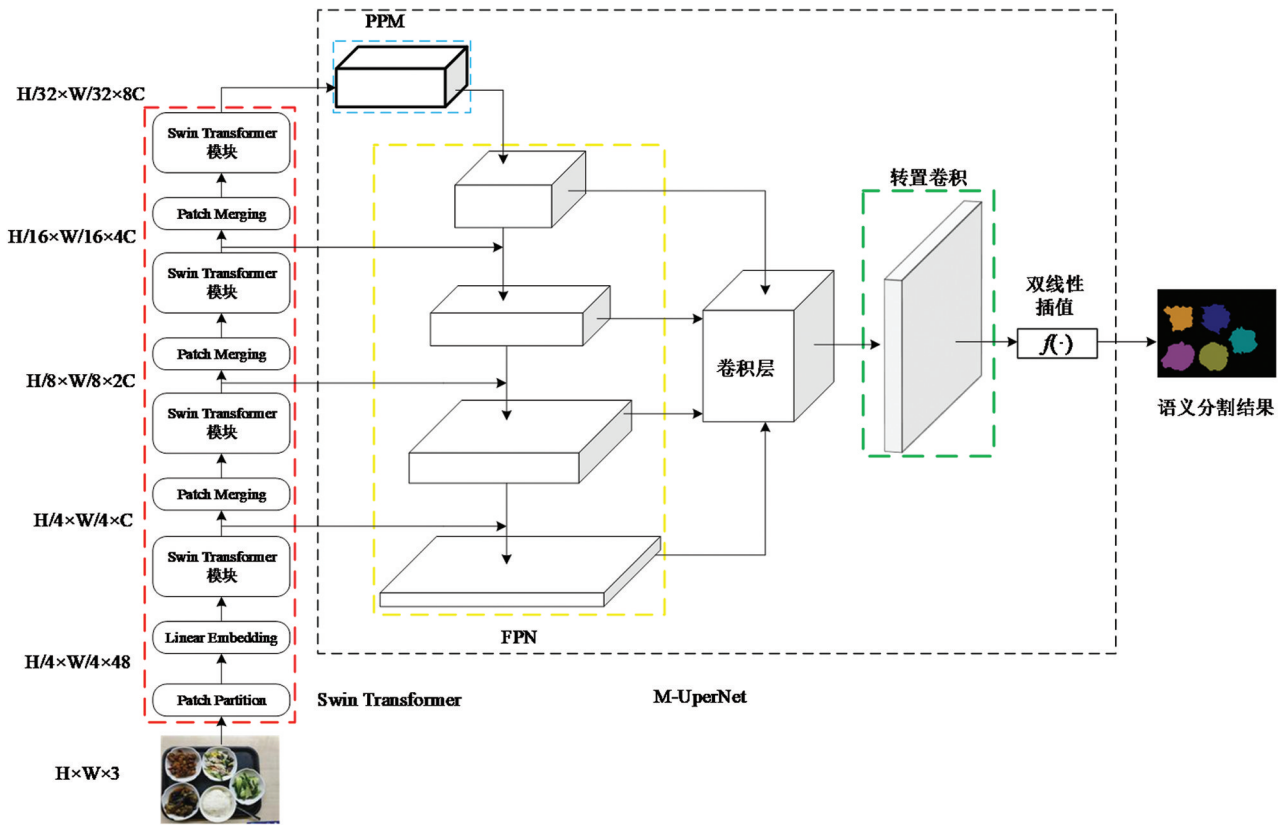


图 6 Swin-T+M-UperNet 模型结构

1.2.2 交叉熵损失函数

假设一张图片的真实标签经过 one-hot 编码后为 Y , 模型预测得到的标签为 X , X 和 Y 的形状为 $H \times$

$W \times C$, H 和 W 分别为图片的高和宽, C 为语义分割的类别数(包括背景), 交叉熵损失^[26]为

$$L = - \sum_{n=1}^N \sum_{c=1}^C y_{n,c} \log \frac{\exp(x_{n,c})}{\sum_{c=1}^C \exp(x_{n,c})}$$

(1)

式中, L 为交叉熵损失, $x_{n,c}$ 为模型预测值 X 中第 n 个像素点标签为 c 的概率, $y_{n,c}$ 为样本真实标签 Y 中第 n 个像素点标签为 c 的概率, 并且 $y_{n,c} \in \{0, 1\}$, N 为像素点总个数, $N = H \times W$.

1.3 浪费度等级定义

假设在一次完整的采样过程中一共采集了 L 幅图像, 某类菜品未被食用(第 1 幅图像)时的像素点分割面积为 S_0 , 食用到第 i 幅图像时的分割面积为 S_i , 用 $\frac{S_i}{S_0}$ 衡量菜品的浪费程度, $\frac{S_i}{S_0}$ 越大说明浪费程度越严重, $\frac{S_i}{S_0} = 0$ 则说明无浪费. 将点 $\left\{ \left(\frac{i}{L-1}, \frac{S_i}{S_0} \right), i = 0, 1, 2, \dots, L-1 \right\}$ 绘制于平面直角坐标系中, 理想的

情况下, 进行均匀采样使得每次减少的面积相同, 则 $0 \leq \frac{S_i}{S_0} \leq 1$, 且 $\frac{S_i}{S_0}$ 呈等差数列递减, 每次的递减量 $d =$

$\frac{1}{L-1}$, 点 $P\left(\frac{p}{L-1}, \frac{S_p}{S_0}\right)$ 和点 $Q\left(\frac{q}{L-1}, \frac{S_q}{S_0}\right)$ 为此过程中的任意两个点, 则两点间的斜率为

$$\frac{\frac{S_p}{S_0} - \frac{S_q}{S_0}}{\frac{p}{L-1} - \frac{q}{L-1}} = \frac{(1-pd) - (1-qd)}{\frac{p}{L-1} - \frac{q}{L-1}} = -1$$

(2)

所以, 理想情况下这些点应该分布在直线 $l_0: y = -x + 1$ 上, 我们将该直线称为理想浪费度变化直线, 然而在实际标注过程中发现由于每次采样的不均匀, 这些点会分布在 l_0 的两侧, 设用最小二乘法拟合这些点得到的实际浪费度变化直线为 $l_1: y = kx + b$. 在理想情况下, 我们将理想浪费度变化直线 l_0 均分为 4 段, 分隔点分别为 $(x_0, y_0) = (0, 1)$, $(x_1, y_1) = (0.25, 0.75)$, $(x_2, y_2) = (0.5, 0.5)$, $(x_3, y_3) = (0.75, 0.25)$, $(x_4, y_4) = (1, 0)$, 并以此划分出 5 个浪费度等级(表 1).

表 1 理想情况下浪费度等级的划分

区间	$y_1 < \frac{S_i}{S_0} \leq y_0$	$y_2 < \frac{S_i}{S_0} \leq y_1$	$y_3 < \frac{S_i}{S_0} \leq y_2$	$y_4 < \frac{S_i}{S_0} \leq y_3$	$\frac{S_i}{S_0} = y_4$
浪费度等级	严重浪费	很浪费	一般浪费	轻度浪费	无浪费

然而, 对于实际浪费度变化直线 l_1 来说, 均匀地划分为 4 段来判断浪费度等级并不是很合理, 直线 l_1 相对于 l_0 来说经过了一定的变换, 如旋转和平移, 因此分隔点也应进行相应的变换, 设变换的映射关系为 F , 则直线 l_1 的分隔点 $(x'_i, y'_i) = F(x_i, y_i)$, $i = 0, 1, 2, 3, 4$. 接下来就求该映射关系 F , 假设 l_0 绕原点旋转 θ 角度(逆时针方向取正, 顺时针方向取负), 并沿 x 轴和 y 轴分别平移 m 和 n 个单位后得到 l_1 (图 7).

A 点为 l_0 的中点坐标, l_0 绕原点旋转 θ 后得到 l_{rotate} , l_{rotate} 经过垂直平移后得到 l_1 , l_{rotate} 上与 A 对应的点为 A' , M 点是 l_0 和 l_{rotate} 的交点, 容易求出 M 点的坐标: $\left(\frac{1}{\text{tg}\left(\alpha + \frac{\theta}{2}\right) + 1}, \frac{\text{tg}\left(\alpha + \frac{\theta}{2}\right)}{\text{tg}\left(\alpha + \frac{\theta}{2}\right) + 1} \right)$, α 为 OA 与横坐标轴的夹角, r 为 OA 的长度. 进而可求出 l_{rotate} 的斜率, 又因为 l_{rotate} 的斜率等于 l_1 的斜率, 于是可由式(3)求出旋转量 θ 的值

$$\frac{r\sin(\alpha+\theta)\left[\operatorname{tg}\left(\alpha+\frac{\theta}{2}\right)+1\right]-\operatorname{tg}\left(\alpha+\frac{\theta}{2}\right)}{r\cos(\alpha+\theta)\left[\operatorname{tg}\left(\alpha+\frac{\theta}{2}\right)+1\right]-1}=k$$

(3)

对于平移量 m 和 n 则可由以下方程组求出

$$\frac{r\sin(\alpha+\theta)+n}{r\cos(\alpha+\theta)+m}=-\frac{1}{k}$$

(4)

$$r\sin(\alpha+\theta)+n=k[r\cos(\alpha+\theta)+m]+b$$

(5)

最后，求出变换后的分隔点

$$\begin{aligned}x_i' &= r_i \cos(\theta + \alpha_i) + m \\ y_i' &= r_i \sin(\theta + \alpha_i) + n \\ r_i &= \sqrt{x_i^2 + y_i^2} \\ \alpha_i &= \arctg\left(\frac{y_i}{x_i}\right)\end{aligned}$$

(6)

式中, $i=1,2,3$, 而 $(x_0, y_0), (x_4, y_4)$ 分别取直线 l_1 与 x 轴和 y 轴的交点.

我们在训练集和验证集的数据上进行了计算, 最终得到浪费度变化的直线分布情况. 图 8a 中的散点为 $\left(\frac{i}{L-1}, \frac{S_i}{S_0}\right)$ 的分布, 图 8b 中理想浪费度变化直线 l_0 上的 5 个点为理想情况下的分隔点, 实际浪费度变化直线 l_1 上的 5 个点为最终计算得到的分隔点(与 l_0 上用相同记号标注的点相对应), 最终, 实际情况下的浪费度等级划分如表 2, 鉴于某些散点存在 $\frac{S_i}{S_0} > y_0$ 的情况, 因此在定义第一个浪费度等级时本文不限制 $\frac{S_i}{S_0}$ 的上界,

其划分阈值为 $\frac{S_i}{S_0} > y_1'$.

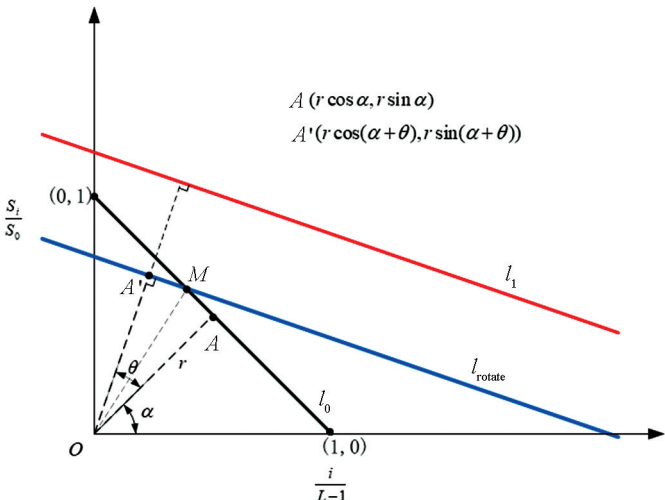
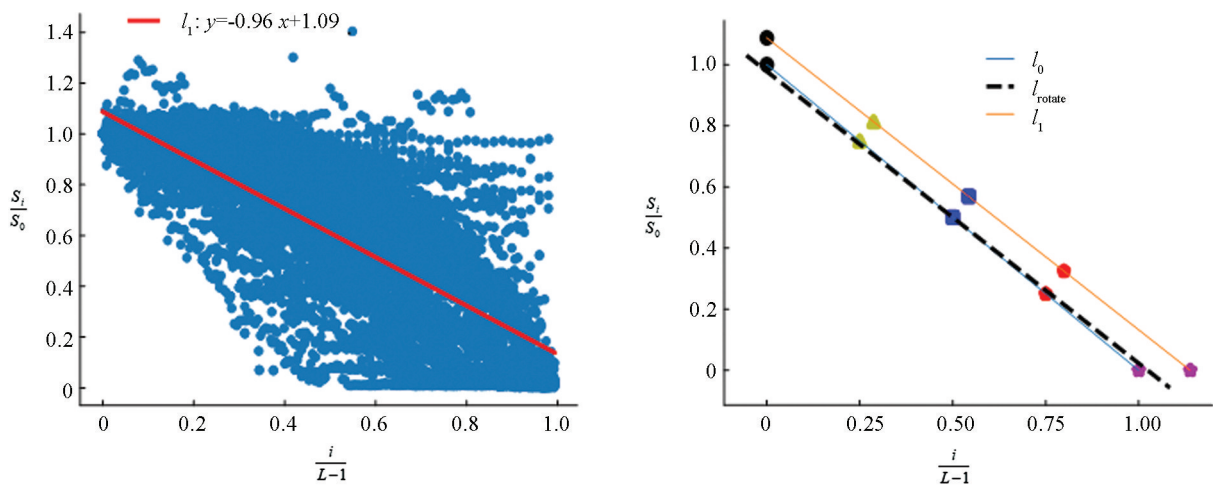


图 7 理想浪费度变化直线与实际浪费度变化直线

布, 图 8b 中理想浪费度变化直线 l_0 上的 5 个点为理想情况下的分隔点, 实际浪费度变化直线 l_1 上的 5 个点为最终计算得到的分隔点(与 l_0 上用相同记号标注的点相对应), 最终, 实际情况下的浪费度等级

划分如表 2, 鉴于某些散点存在 $\frac{S_i}{S_0} > y_0$ 的情况, 因此在定义第一个浪费度等级时本文不限制 $\frac{S_i}{S_0}$ 的上界,

其划分阈值为 $\frac{S_i}{S_0} > y_1'$.



(a) 散点图与回归直线 (b) 各直线间的变化关系

图 8 浪费度变化直线

表 2 实际情况下浪费度等级的划分

区间	$y_1'=0.81<\frac{S_i}{S_0}$	$y_2'=0.56<\frac{S_i}{S_0}\leq y_1'=0.81$	$y_3'=0.32<\frac{S_i}{S_0}\leq y_2'=0.56$	$y_4'=0<\frac{S_i}{S_0}\leq y_3'=0.32$	$\frac{S_i}{S_0}=0=y_4'$
浪费度等级	严重浪费	很浪费	一般浪费	轻度浪费	无浪费

2 结果与分析

2.1 模型训练结果

模型训练的环境为 Python: 3.7, Pytorch: 1.11.0, CUDA: 11.3, GPU: RTX3060Ti; 训练过程中配置的超参数为输入图像大小: 224×224 , 批量大小: 24, 学习率(lr): 按 $lr=lr_0\left(1-\frac{e}{E}\right)^{0.9}$ 进行衰减, lr_0 表示初始学习率, 设置为 $5e^{-4}$, E 表示最大迭代次数, 设置为 200, e 表示当前迭代次数. 迭代过程中, 若验证集的平均交并比在 20 个回合内不再上升就停止迭代.

本文中用到的分割效果评价指标有平均像素精度(m_{AP}), 平均交并比(m_{IOU})以及 F_1 精度^[27], 用 p_{ii} 表示真正例, p_{ij} 和 p_{ji} 分别表示假正例和假负例, P 和 R 分别表示精度和召回率, 各评价指标计算式为

$$m_{AP}=\frac{1}{C}\sum_{i=1}^C\frac{p_{ii}}{\sum_{j=1}^Cp_{ij}}$$

(7)

$$m_{IOU}=\frac{1}{C}\sum_{i=1}^C\frac{p_{ii}}{\sum_{j=1}^Cp_{ij}+\sum_{j=1}^Cp_{ji}-p_{ii}}$$

(8)

$$F_1=\frac{2PR}{P+R}$$

(9)

不同模型在测试集上的评价指标值如表 3. 从表 3 可以看出, 相较于以卷积神经网络作为特征提取结构的模型 DeepLabV3+^[14], PSPNet^[28], 以 Swin-Transformer 作为特征提取结构的模型分割效果会更好. 此外, 本文提出的 Swin-T+M-UperNet 模型效果更优, 其平均交并比(m_{IOU})较原模型提高了 1.02%. 在模型的解码器部分, 通过引入转置卷积, 增加了可学习的参数, 相较于直接通过双线性插值恢复出和输入图像相同大小的掩码, 可以保留更多的细节信息.

表 3 模型在测试集上的语义分割评价指标

模型	$m_{AP}/\%$	$m_{IOU}/\%$	$F_1/\%$
DeepLabV3+	95.92	91.30	95.37
PSPNet	94.43	89.63	94.46
Swin-T+UperNet	95.66	92.28	95.96
Swin-T+M-UperNet	96.72	93.30	96.51

2.2 浪费度检测

实验结果表明, 菜品图像的语义分割结果中 Swin-T+M-UperNet 模型的分割效果最好, 因此本文选用该模型进行浪费度检测, 流程如图 9. 将未食用时的菜品图像和食用后的菜品图像分别输入到训练好的 Swin-T+M-UperNet 模型中得到对应的分割掩码, 然后计算对应菜品分割掩码区域面积(像素点个数)的比值来判断每一类菜品的浪费度等级, 右侧检测结果中的检测框为菜品未食用时每一个菜品分割区域的最大外接矩形框. 本文通过该区域来确定菜品的种类和位置, 如图 9 中初始时刻图像得到的分割掩码有 6 类:

5 类菜品加上 1 类背景(黑色区域), 图中红色掩码区域对应菜品种类 9, 取该最大连通区域的外接矩形即可确定该类菜品位置, 并可在矩形框外标注出菜品种类以及浪费度等级。

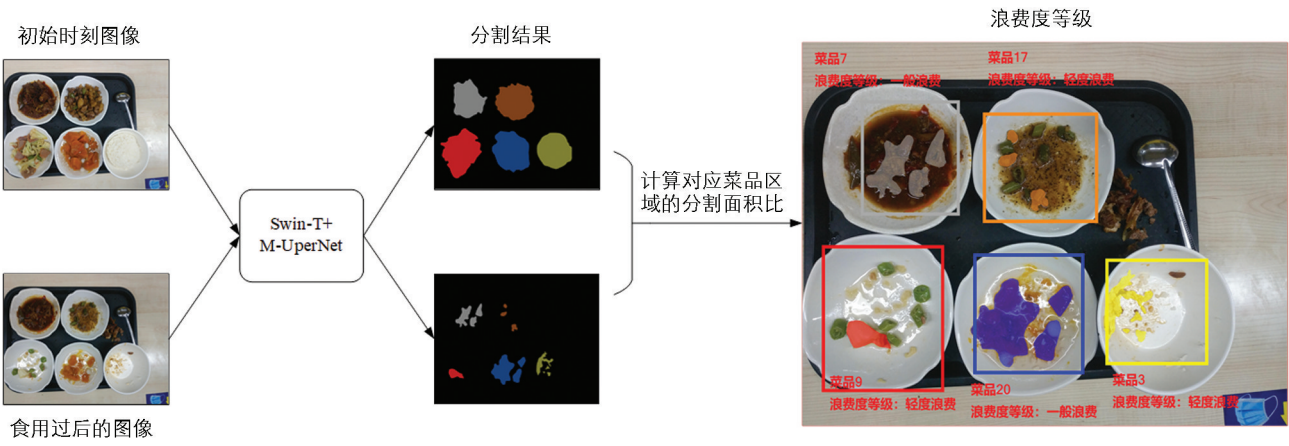


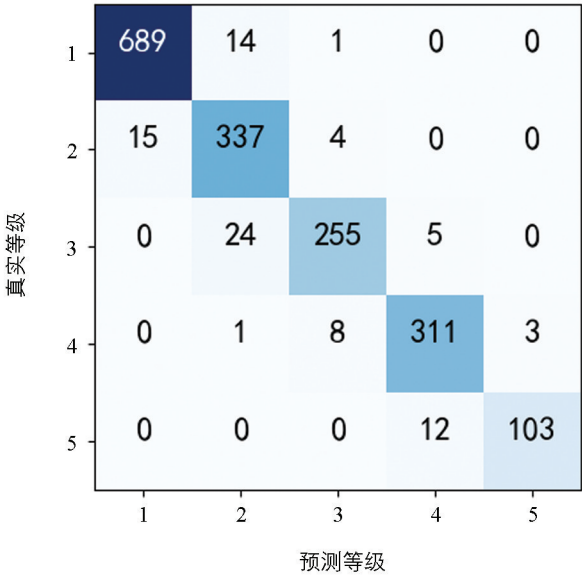
图 9 浪费度检测流程

按照浪费度等级的定义, 模型对测试集上的 1 782 个菜品样本进行了检测, 最终检测结果的混淆矩阵分布情况如图 10. 从混淆矩阵可以看出, 错判的样本主要分布在主对角线两侧, 即这些样本的真实浪费度等级和错判等级之间是相邻的, 检测的整体准确率为 $(689 + 337 + 255 + 311 + 103) / 1\,782 = 95.12\%$.

3 结论

本文将图像语义分割算法应用于菜品浪费度检测中, 在 Swin Transformer 和 UperNet 结构的基础上进行改进, 增加了一层转置卷积, 并传统的以卷积神经网络为特征提取结构的模型进行了比较. 在测试集上的分割结果表明, 以 Transformer 为特征提取结构的模型分割效果较好, 改进的 Swin-T+M-UperNet 模型相较于原来的 Swin-T+UperNet 模型分割效果更优, 在测试集上的平均交并比可以达到 93.30%, 提升了 1.02%.

本文还根据训练集和验证集上菜品食用前后分割面积比的统计信息制定出了浪费度等级, 首先在理想直线上划分出浪费度等级的分隔点, 然后求出理想直线到实际直线间的变换关系, 根据该变换关系求出实际直线上对应的分隔点, 按照分隔点的划分阈值, 将菜品的浪费程度划分为严重浪费、很浪费、一般浪费、轻度浪费和无浪费 5 个等级, 同时提出了菜品浪费度检测的流程和方法, 最终在测试集上的浪费度等级识别准确率能达到 95.12%. 本文提出的检测方法具有一定的合理性, 为餐饮浪费的检测和管理提供了一种可行的技术方法, 可以帮助餐厅监测菜品的浪费程度, 并从菜单里淘汰掉一些浪费程度较高的菜品, 保留浪费度较低的菜品, 从而做到从源头上减少浪费, 也可以帮助优化餐馆的进货清单, 提升餐馆的经济效益.



1 为严重浪费, 2 为很浪费, 3 为一般浪费, 4 为轻度浪费, 5 为无浪费.

图 10 混淆矩阵

本文所研究的菜品对象在食用过程中像素分割面积呈现递减的趋势,但对于一些在食用过程中质量减少而像素分割面积变化很小的菜品(如汤类菜品),该方法还有一定的局限性,对此类菜品的检测还有待进一步去探索.

参考文献:

- [1] 李贺,李鸣晓,孟繁华,等. 我国餐厅食物浪费现状调查及影响因素分析 [J]. 环境工程技术学报, 2021, 11(5): 898-907.
- [2] WANG L E, LIU G, LIU X, et al. The Weight of Unfinished Plate: a Survey Based Characterization of Restaurant Food Waste in Chinese Cities [J]. Waste Management, 2017, 66: 3-12.
- [3] YANG Y, BAO W, XIE G H. Estimate of Restaurant Food Waste and Its Biogas Production Potential in China [J]. Journal of Cleaner Production, 2019, 211: 309-320.
- [4] CHEN H, JIANG W, YANG Y, et al. State of the Art on Food Waste Research: a Bibliometrics Study from 1997 to 2014 [J]. Journal of Cleaner Production, 2017, 140: 840-846.
- [5] LANG L, WANG Y, CHEN X, et al. Awareness of Food Waste Recycling in Restaurants: Evidence from China [J]. Resources, Conservation and Recycling, 2020, 161: 104949.
- [6] CHEN J J, NGO C W. Deep-Based Ingredient Recognition for Cooking Recipe Retrieval [C] //Proceedings of the 24th ACM international conference on Multimedia. Amsterdam, The Netherlands. New York: ACM, 2016: 32-41.
- [7] 王晓朋. 精细粒度的菜品识别方法研究 [D]. 成都: 电子科技大学, 2020.
- [8] 吴正东. 基于深度学习的中餐菜品图像分类算法研究 [D]. 成都: 电子科技大学, 2020.
- [9] 何杰. 基于集成深度卷积神经网络的中餐菜品识别 [D]. 广州: 华南理工大学, 2018.
- [10] EGE T, YANAI K. Simultaneous Estimation of Dish Locations and Calories with Multi-Task Learning [J]. IEICE Transactions on Information and Systems, 2019(7): 1240-1246.
- [11] 姚华莹, 彭亚雄, 陆安江. 基于轻量化 SSD 的菜品识别 [J]. 智能计算机与应用, 2021, 11(8): 23-30.
- [12] POPLY P, ARUL JOTHI J A. Refined Image Segmentation for Calorie Estimation of Multiple-Dish Food Items [C] // 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS). Greater Noida, India. IEEE, 2021: 682-687.
- [13] OKAMOTO K, ADACHI K, YANAI K. Region-Based Food Calorie Estimation for Multiple-Dish Meals [C] //Proceedings of the 13th International Workshop on Multimedia for Cooking and Eating Activities. New York: ACM, 2021: 17-24.
- [14] YANG Z, PENG X B, YIN Z J, et al. Deeplab_v3_plus-Net for Image Semantic Segmentation with Channel Compression [C] //2020 IEEE 20th International Conference on Communication Technology (ICCT). Nanning, China. IEEE, 2020: 1320-1324.
- [15] HIPPOCRATE E A A, SUWA H, ARAKAWA Y, et al. Food Weight Estimation Using Smartphone and Cutlery [C] // Proceedings of the First Workshop on IoT-enabled Healthcare and Wellness Technologies and Systems. Singapore, Singapore. New York: ACM, 2016: 9-14.
- [16] CHEN M Y, YANG Y H, HO C J, et al. Automatic ChineseFood Identification and Quantity Estimation [C] //SA '12: SIGGRAPH Asia 2012 Technical Briefs. Singapore, Singapore. New York: ACM, 2012: 1-4.
- [17] LOWE D G. DistinctiveImage Features from Scale-Invariant Keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [18] HU X H, YUAN D, SUN M G, et al. A Slice-Based Method for Food Volume Estimation [C] //Proceedings of the

2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence. Sanya, China. New York: ACM, 2019: 1-6.

[19] SHELHAMER E, LONG J, DARRELL T. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.

[20] XU B B, YANG F, YANG J F, et al. SPNet: Superpixel Pyramid Network for Scene Parsing [C] //2018 Chinese Automation Congress (CAC). Xi'an, China. IEEE, 2019: 3690-3695.

[21] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C] //2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy. IEEE, 2017: 2980-2988.

[22] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada. IEEE, 2022: 9992-10002.

[23] ZHENG W B, YAN L, WANG F Y, et al. Progressive Knowledge-Embedded Unified Perceptual Parsing for Scene Understanding [C] //2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville, TN, USA. IEEE, 2021: 1633-1642.

[24] 石洪康, 田涯涯, 杨创, 等. 基于卷积神经网络的家蚕幼虫品种智能识别研究 [J]. 西南大学学报(自然科学版), 2020, 42(12): 34-45.

[25] ZEILER M D, TAYLOR G W, FERGUS R. AdaptiveDeconvolutional Networks for Mid and High Level Feature Learning [C] //2011 International Conference on Computer Vision. Barcelona, Spain. IEEE, 2012: 2018-2025.

[26] JADON S. ASurvey of Loss Functions for Semantic Segmentation [C] //2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). Via del Mar, Chile. IEEE, 2020: 1-7.

[27] MINAE S, BOYKOV Y, PORIKLI F, et al. Image Segmentation Using Deep Learning: a Survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(7): 3523-3542.

[28] ZHAO H S, SHI J P, QI X J, et al. Pyramid Scene Parsing Network [C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. IEEE, 2017: 6230-6239.

责任编辑 周仁惠